

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 20:06:21 ; Search time 46.4249 Seconds
(without alignments)
875.264 Million cell updates/sec

Title: US-09-677-374-6

Perfect score: 1335

Sequence: 1 MSVEFYNSKSAQTSITPI.....IYGTAQPQDGRWQVISTEK 256

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A. Geneseq 19Jun03.*

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1335	100.0	256	22	AAE05747
2	1335	100.0	256	23	AAU97869
3	832	62.3	161	22	AAE05748
4	832	62.3	161	23	AAU97868
5	815	61.0	162	22	AAE05749
6	815	61.0	162	23	AAU97867
7	815	61.0	162	23	AAU97867
8	487	36.5	190	22	AAE05746
9	487	36.5	328	22	AAE05749

10	487	36.5	341	22	AAE05747	Clostridium cellul
11	483	36.2	162	15	AAE05747	Cellulose binding
12	483	36.2	163	22	AAE05745	Clostridium cellul
13	483	36.2	428	22	AAE05748	Clostridium cellul
14	480.5	36.0	382	20	AAE05748	Gaussia luciferase
15	480.5	36.0	382	23	AAE13383	Gaussia species CB
16	477	35.7	162	20	AAW90077	C. cellulovorans C
17	472	35.4	154	20	AAW90081	C. cellulovorans C
18	472	35.4	156	20	AAW90080	C. cellulovorans C
19	245	18.4	167	17	AAE05080	Cellulose binding
20	245	18.4	476	21	AAE05080	A mannanase-linker
21	245	18.4	493	20	AAE05080	Pectate lyase-link
22	245	18.4	493	20	AAE05080	Pectate lyase CBD
23	245	18.4	531	18	AAW15238	Scaffoldin protein
24	245	18.4	1853	19	AAW43108	C. thermocellum ce
25	149.5	11.2	1426	20	AAE13492	Truncated cellulase
26	149.5	11.2	1426	23	AAE13492	Active cellulase p
27	148.5	11.1	616	20	AAE13494	Truncated cellulase
28	148.5	11.1	616	23	AAE13494	Active cellulase h
29	148.5	11.1	1751	20	AAE13493	Truncated cellulase
30	148.5	11.1	1751	23	AAE13493	Active cellulase p
31	127.5	9.6	182	24	ABJ18820	Pseudomonas aerugi
32	127.5	9.6	1352	22	AAE05747	Amino acid sequenc
33	126.5	9.5	1350	22	AAE05747	Amino acid sequenc
34	124.5	9.3	153	24	ABP71658	A. cellulolyticus
35	124.5	9.3	1121	24	ABP71656	A. cellulolyticus
36	122.5	9.2	551	18	AAW18790	Corrected Bacillus
37	121	9.1	150	24	ABP73025	Amino acid sequenc
38	121	9.1	1228	24	ABP73025	Amino acid sequenc
39	120.5	9.0	154	24	ABP73020	Amino acid sequenc
40	120.5	9.0	762	24	ABP73022	Amino acid sequenc
41	115	8.6	88	24	ABP73018	Amino acid sequenc
42	115	8.6	89	24	ABP73017	Amino acid sequenc
43	115	8.6	477	23	AAO22444	Protein encoded by
44	115	8.6	496	23	AAO22443	Protein encoded by
45	115	8.6	677	23	AAU98063	Bacillus subtilis

ALIGNMENTS

RESULT 1

AAE05747

ID AAE05747 standard; Protein; 256 AA.

AC AAE05747;

XX 11-JUL-2001 (first entry)

DT C17E2 OspA construct with N-terminal fusion partner.

DE Podkilothermic fish; Piscirickettsia salmonis; rickettsial pathogen;
KW vaccine; OspA; salmonid rickettsial septicaemia; rickettsial disease;
KW SRS; 17E2; fusion construct.

XX Piscirickettsia salmonis.

OS Synthetic.

XX Key Location/Qualifiers

FT Region 1..95

FT Region /label= Undefined_N-terminal_fusion_partner

FT Region 96..256

FT Region /label= C17E2 OspA

FT Region /note= "Product of OspA gene optimised for expression in Escherichia coli"

XX (KAYW/) KAY W W.
 PA (BURI/) BURIAN J.
 PA (KUZU/) KUZIK M A.
 XX
 XX Kay WW, Burian J, Kuzyk MA;
 PI
 XX WPI; 2001-316944/34.
 DR N-PSDB; AAF86248.
 XX
 XX Method for protecting poikilothermic fish against salmonid rickettsial
 PT septicaemia and other rickettsial diseases comprises administering a
 PT vaccine containing the OspA protein of *Piscirickettsia salmonis*.
 XX
 XX Example 4; Fig 5; 35pp; English.
 XX
 XX This invention relates to a method for the protection against infection
 CC of a poikilothermic fish by the bacterial pathogen, *Piscirickettsia*
 CC salmonis. The method comprises administering an immunogenic amount of a
 CC P. salmonis specific antigen termed OspA, or an immunogenic fragment of
 CC OspA in the form of a vaccine. The method is used for protecting animals,
 CC particularly poikilothermic fish, against the bacterial pathogen
 CC P. salmonis. The method is also useful for protecting against salmonid
 CC rickettsial septicaemia (SRS) and other rickettsial diseases. The present
 CC sequence represents the amino acid sequence of C17E2, a P. salmonis OspA
 CC construct optimised for expression in *Escherichia coli*, fused to an
 CC undefined N-terminal fusion partner. The fusion protein is used in a
 CC vaccine to create an anti-OspA antibody response.
 XX
 XX Sequence 256 AA;

Query Match 100.0%; Score 1335; DB 22; Length 256;
 Best Local Similarity 100.0%; Pred. No. 9.1e-123;
 Matches 256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSVEFYNSKSAQNSITPIIKITNTSDSLNLDVVKVRYVYTSQTQGTQWCDHAGAL 60
 DB 1 MSVEFYNSKSAQNSITPIIKITNTSDSLNLDVVKVRYVYTSQTQGTQWCDHAGAL 60
 QY 61 LGNSYVDNTSKVTANFVKETASPTSTYDYLDP SHMRGCLQGSSLIISVFLVGCANFS 120
 DB 61 LGNSYVDNTSKVTANFVKETASPTSTYDYLDP SHMRGCLQGSSLIISVFLVGCANFS 120
 QY 121 ROEVGAATGAVVGGVAGQLFGKSGRVSMAIGGAVLGLGSKIGOSMDQDKIKLNQSL 180
 DB 121 ROEVGAATGAVVGGVAGQLFGKSGRVSMAIGGAVLGLGSKIGOSMDQDKIKLNQSL 180
 QY 181 EKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKKQERRQYCREFOQKAMIAGQKEIYGT 240
 DB 181 EKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKKQERRQYCREFOQKAMIAGQKEIYGT 240
 QY 241 ACPQPDGRWQVISTEK 256
 DB 241 ACPQPDGRWQVISTEK 256

RESULT 2
 AAU97869
 ID AAU97869 standard; Protein; 256 AA.
 XX
 XX AAU97869;
 XX
 XX 12-AUG-2002 (first entry)
 XX
 XX E. coli codon optimised OspA, 17e2 with N-terminal fusion peptide.
 XX
 XX Outer surface lipoprotein; OspA; antibacterial; immunosuppressive;
 KW vaccine; poikilothermic fish; fin-fish; rickettsial septicaemia;
 KW Rickettsial disease; 17e2.
 XX
 XX *Piscirickettsia salmonis*.
 OS
 OS Synthetic.
 XX

PN CA2339327-A1.
 XX
 PD 15-MAR-2002.
 XX
 XX 19-MAR-2001; 2001CA-2339327.
 PF
 XX 15-SEP-2000; 2000US-0677374.
 PR
 XX (THOR/) THORNTON J C.
 PA (KAYW/) KAY W W.
 PA (BURI/) BURIAN J.
 PA (KUZU/) KUZIK M A.
 XX
 PI Thornton JC, Kay WW, Burian J, Kuzyk MA;
 XX
 XX WPI; 2002-455221/49.
 DR N-PSDB; ABK52403.
 DR
 XX Inducing immunity in fin fish to Rickettsial septicaemia, comprises
 PT administration of an outer surface lipoprotein (OspA) of a bacterial
 PT strain, as a vaccine -
 XX
 XX Example 4; Fig 5; 55pp; English.
 XX
 XX The invention describes a method of protecting a poikilothermic fish
 CC against infection by the bacterial pathogen *Piscirickettsia salmonis*
 CC comprising administering either intraperitoneally, by immersion or
 CC orally, an immunogenic amount of principal antigen, the OspA (outer
 CC surface lipoprotein), its variants, non-lipidated form or antigenic
 CC peptides derived or synthesized with or without an adjuvant. The new
 CC method is used to provide an outer surface lipoprotein (OspA) of
 CC bacterial strain *Piscirickettsia salmonis* as a vaccine to induce immunity
 CC in fin-fish against Rickettsial septicaemia and other related
 CC Rickettsial diseases caused by either a virus, bacteria or parasite.
 CC This is the amino acid sequence of the *Escherichia coli* codon optimised
 CC outer surface lipoprotein OspA (17e2) with an N-terminal fusion used in
 CC the creation of the vaccine described in the invention.
 XX
 XX Sequence 256 AA;

Query Match 100.0%; Score 1335; DB 23; Length 256;
 Best Local Similarity 100.0%; Pred. No. 9.1e-123;
 Matches 256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSVEFYNSKSAQNSITPIIKITNTSDSLNLDVVKVRYVYTSQTQGTQWCDHAGAL 60
 DB 1 MSVEFYNSKSAQNSITPIIKITNTSDSLNLDVVKVRYVYTSQTQGTQWCDHAGAL 60
 QY 61 LGNSYVDNTSKVTANFVKETASPTSTYDYLDP SHMRGCLQGSSLIISVFLVGCANFS 120
 DB 61 LGNSYVDNTSKVTANFVKETASPTSTYDYLDP SHMRGCLQGSSLIISVFLVGCANFS 120
 QY 121 ROEVGAATGAVVGGVAGQLFGKSGRVSMAIGGAVLGLGSKIGOSMDQDKIKLNQSL 180
 DB 121 ROEVGAATGAVVGGVAGQLFGKSGRVSMAIGGAVLGLGSKIGOSMDQDKIKLNQSL 180
 QY 181 EKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKKQERRQYCREFOQKAMIAGQKEIYGT 240
 DB 181 EKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKKQERRQYCREFOQKAMIAGQKEIYGT 240
 QY 241 ACPQPDGRWQVISTEK 256
 DB 241 ACPQPDGRWQVISTEK 256

RESULT 3
 AAB81127
 ID AAB81127 standard; Protein; 161 AA.
 XX
 XX AAB81127;
 XX
 XX 11-JUL-2001 (first entry)
 XX
 XX

DE Optimised Ospa protein 17E2 amino acid sequence.

XX Poikilothermic fish; Piscirickettsia salmonis; rickettsial pathogen;

KW vaccine; Ospa; salmonid rickettsial septicaemia; rickettsial disease;

KW SRS.

XX Piscirickettsia salmonis.

OS Synthetic.

XX Key Location/Qualifiers

FT 109..128

FT /label= B_cell_epitope

XX CA2281913-A1.

XX 17-MAR-2001.

XX 17-SEP-1999; 99CA-2281913.

XX 17-SEP-1999; 99CA-2281913.

XX (KAYW/) KAY W W.

PA (BURI/) BURIAN J.

PA (KUZY/) KUZYSK M A.

XX Kay WW, Burian J, Kuzyk MA;

XX WPI; 2001-316844/34.

DR N-PSDB; AAF86247.

XX Method for protecting poikilothermic fish against salmonid rickettsial

PT septicaemia and other rickettsial diseases comprises administering a

PT vaccine containing the Ospa protein of Piscirickettsia salmonis

XX Disclosure; Fig 5; 35pp; English.

XX This invention relates to a method for the protection against infection

CC of a poikilothermic fish by the bacterial pathogen, Piscirickettsia

CC salmonis. The method comprises administering an immunogenic amount of a

CC P. salmonis specific antigen termed Ospa, or an immunogenic fragment of

CC Ospa in the form of a vaccine. The method is used for protecting animals,

CC particularly poikilothermic fish, against the bacterial pathogen

CC P. salmonis. The method is also useful for protecting against salmonid

CC rickettsial septicaemia (SRS) and other rickettsial diseases. The present

CC sequence represents optimised P. salmonis Ospa protein 17E2. The DNA

CC encoding Ospa 17E2 (AAF86247) has been optimised for expression in

CC Escherichia coli. An Ospa protein with an N-terminal fusion partner is

CC used in a vaccine to create an anti-Ospa antibody response.

XX Sequence 161 AA;

Query Match 62.3%; Score 832; DB 22; Length 161;
Best Local Similarity 100.0%; Pred. No. 1.4e-73;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 MRGCLQSSLLIIISVFLVCCQNFQEVGAATGAVGVGAGQLFGKSGRYSMAIGGAV 155
DB 1 MRGCLQSSLLIIISVFLVCCQNFQEVGAATGAVGVGAGQLFGKSGRYSMAIGGAV 60

QY 156 LGGLIGSKIGSQMDQDDKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPRTYQRYNKE 215
DB 61 LGGLIGSKIGSQMDQDDKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPRTYQRYNKE 120

QY 216 RQOYCRFPQKAMTAGQKQEIYGTACPDGRWQVISTEK 256
DB 121 RQOYCRFPQKAMTAGQKQEIYGTACPDGRWQVISTEK 161

RESULT 4
AAU97868
ID AAU97868 standard; Protein; 161 AA.
XX
AC AAU97868;

XX 12-AUG-2002 (first entry)

DT Escherichia coli codon optimised Ospa, 17e2.

DE

XX Outer surface lipoprotein; Ospa; antibacterial; immunosuppressive;

KW vaccine; poikilothermic fish; fin-fish; Rickettsial septicaemia;

KW Rickettsial disease; 17e2.

XX Piscirickettsia salmonis.

OS Synthetic.

XX CA2339327-A1.

XX 15-MAR-2002.

XX 19-MAR-2001; 2001CA-2339327.

XX 15-SEP-2000; 2000US-0677374.

XX (THOR/) THORNTON J C.

PA (KAYW/) KAY W W.

PA (BURI/) BURIAN J.

PA (KUZY/) KUZYSK M A.

XX Thornton JC, Kay WW, Burian J, Kuzyk MA;

XX WPI; 2002-455221/49.

DR N-PSDB; ABK52402.

XX Inducing immunity in fin fish to Rickettsial septicaemia, comprises

PT administration of an outer surface lipoprotein (Ospa) of a bacterial

PT strain, as a vaccine

XX Example 4; Fig 5; 55pp; English.

XX The invention describes a method of protecting a poikilothermic fish

CC against infection by the bacterial pathogen Piscirickettsia Salmonis

CC comprising administering either intraperitoneally, by immersion or

CC orally, an immunogenic amount of principal antigen, the Ospa (outer

CC surface lipoprotein), its variants, non-lipidated form or antigenic

CC peptides derived or synthesized with or without an adjuvant. The new

CC method is used to provide an outer surface lipoprotein (Ospa) of

CC bacterial strain Piscirickettsia salmonis as a vaccine to induce immunity

CC in fin-fish against Rickettsial septicaemia and other related

CC Rickettsial diseases caused by either a virus, bacteria or parasite.

CC This is the amino acid sequence of the Escherichia coli codon optimised

CC outer surface lipoprotein Ospa (17e2) used in the creation of the vaccine

CC described in the invention.

XX Sequence 161 AA;

Query Match 62.3%; Score 832; DB 23; Length 161;
Best Local Similarity 100.0%; Pred. No. 1.4e-73;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 MRGCLQSSLLIIISVFLVCCQNFQEVGAATGAVGVGAGQLFGKSGRYSMAIGGAV 155
DB 1 MRGCLQSSLLIIISVFLVCCQNFQEVGAATGAVGVGAGQLFGKSGRYSMAIGGAV 60

QY 156 LGGLIGSKIGSQMDQDDKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPRTYQRYNKE 215
DB 61 LGGLIGSKIGSQMDQDDKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPRTYQRYNKE 120

QY 216 RQOYCRFPQKAMTAGQKQEIYGTACPDGRWQVISTEK 256
DB 121 RQOYCRFPQKAMTAGQKQEIYGTACPDGRWQVISTEK 161

RESULT 5
AAG78025
ID AAG78025 standard; Protein; 162 AA.
XX

AC AAG78025;
 XX
 DT 15-JAN-2002 (first entry)
 XX
 DE Piscirickettsia salmonis polypeptide P10.6.
 XX
 DE Piscirickettsia salmonis; piscirickettsiosis; salmonid rickettsial;
 KW septicaemia; SRS; surface antigen; vaccine; antibacterial; fish;
 KW ATCC VR-1361.
 XX
 OS Piscirickettsia salmonis.
 XX
 PN WO200168865-A2.
 XX
 PD 20-SEP-2001.
 XX
 PF 12-MAR-2001; 2001WO-GB01055.
 XX
 PR 11-MAR-2000; 2000GB-0005838.
 PR 01-JUL-2000; 2000GB-0016080.
 PR 01-JUL-2000; 2000GB-0016082.
 PR 29-JUL-2000; 2000GB-0018599.
 XX
 PA (AQUA-) AQUA HEALTH EURO LTD.
 XX
 XX Simard N, Brouwers H, Jones S, Griffiths S, Valenzuela P;
 PI Burzio L;
 XX
 XX WPI; 2001-639050/73.
 DR N-PSDB; AAH79040.
 XX
 XX New nucleic acids encoding an amino acid sequence homologous to the
 PT surface antigen present on Piscirickettsia salmonis are useful to
 PT protect fish against piscirickettsiosis -
 PT
 XX Claim 6; Fig 5; 25pp; English.
 XX
 CC The invention relates to nucleic acid sequences and the encoded protein
 CC of a least part of the surface antigen present on Piscirickettsia
 CC salmonis for production of a vaccine with antibacterial activity to
 CC protect fish against P. salmonis which causes piscirickettsiosis, also
 CC known as salmonid rickettsial septicaemia.
 CC
 XX Sequence 162 AA;
 SQ
 Query Match 61.0%; Score 815; DB 22; Length 162;
 Best Local Similarity 98.8%; Pred. No. 6.4e-72;
 Matches 158; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 97 RGCLOGSSLIISVFLVGCQNFSRQEVGAATGAVVGVAGQLFGKSGRVSMAIGGAVL 156
 Db 3 RGCLOGSSLIISVFLVGCQNFSRQEVGAATGAVVGVAGQLFGKSGRVSMAIGGAVL 62
 QY 157 GGLIGSKIQSQMDQDKIKLNQSLKVKAGQVTRWNPDTGNSYSVEPVRTYQRYNKQER 216
 Db 63 GGLIGSKIQSQMDQDKIKLNQSLKVKAGQVTRWNPDTGNSYSVEPVRTYQRYNKQER 122
 QY 217 RQCYCREFOQKAMIAQKQEIYGTACPDGRWQVISTEK 256
 Db 123 RQCYCREFOQKAMIAQKQEIYGTACPDGRWQVISTEK 162
 RESULT 6
 AAB81126
 ID AAB81126 standard; Protein; 162 AA.
 XX
 AC AAB81126;
 XX
 DT 11-JUL-2001 (first entry)
 XX
 DE OspA antigen amino acid sequence.
 XX
 XX Poikilothermic fish; Piscirickettsia salmonis; rickettsial pathogen;

KW vaccine; OspA; salmonid rickettsial septicaemia; rickettsial disease;
 KW SRS.
 XX
 OS Piscirickettsia salmonis.
 XX
 FH Key Location/Qualifiers
 FT Region 110..129
 XX /label= B_cell_epitope
 PN CA2281913-A1.
 XX
 PD 17-MAR-2001.
 XX
 PF 17-SEP-1999; 99CA-2281913.
 XX
 PR 17-SEP-1999; 99CA-2281913.
 XX
 PA (KAYW/) KAY W W.
 PA (BURI/) BURIAN J.
 PA (KUZU/) KUZUK M A.
 XX
 PI Kay WW, Burian J, Kuzyk MA;
 XX
 DR WPI; 2001-316944/34.
 DR N-PSDB; AAF86246.
 XX
 PT Method for protecting poikilothermic fish against salmonid rickettsial
 PT septicaemia and other rickettsial diseases comprises administering a
 PT vaccine containing the OspA protein of Piscirickettsia salmonis -
 XX
 PS Example 2; Fig 2B; 35pp; English.
 XX
 CC This invention relates to a method for the protection against infection
 CC of a poikilothermic fish by the bacterial pathogen, Piscirickettsia
 CC salmonis. The method comprises administering an immunogenic amount of a
 CC P. salmonis specific antigen termed OspA, or an immunogenic fragment of
 CC OspA in the form of a vaccine. The method is used for protecting animals,
 CC particularly poikilothermic fish, against the bacterial pathogen
 CC P. salmonis. The method is also useful for protecting against salmonid
 CC rickettsial septicaemia (SRS) and other rickettsial diseases. The present
 CC sequence represents P. salmonis OspA protein. An OspA protein with an
 CC N-terminal fusion partner is used in a vaccine to create an anti-OspA
 CC antibody response.
 XX
 SQ Sequence 162 AA;
 Query Match 61.0%; Score 815; DB 22; Length 162;
 Best Local Similarity 98.8%; Pred. No. 6.4e-72;
 Matches 158; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 97 RGCLOGSSLIISVFLVGCQNFSRQEVGAATGAVVGVAGQLFGKSGRVSMAIGGAVL 156
 Db 3 RGCLOGSSLIISVFLVGCQNFSRQEVGAATGAVVGVAGQLFGKSGRVSMAIGGAVL 62
 QY 157 GGLIGSKIQSQMDQDKIKLNQSLKVKAGQVTRWNPDTGNSYSVEPVRTYQRYNKQER 216
 Db 63 GGLIGSKIQSQMDQDKIKLNQSLKVKAGQVTRWNPDTGNSYSVEPVRTYQRYNKQER 122
 QY 217 RQCYCREFOQKAMIAQKQEIYGTACPDGRWQVISTEK 256
 Db 123 RQCYCREFOQKAMIAQKQEIYGTACPDGRWQVISTEK 162
 RESULT 7
 AAU97867
 ID AAU97867 standard; Protein; 162 AA.
 XX
 AC AAU97867;
 XX
 DT 12-AUG-2002 (first entry)
 XX
 DE Piscirickettsia salmonis outer surface lipoprotein OspA.
 XX

KW Outer surface lipoprotein; OspA; antibacterial; immunosuppressive;
 KW vaccine; poikilothermic fish; fin-fish; Rickettsial septicaemia;
 KW Rickettsial disease.
 XX Piscirickettsia salmonis.
 XX CA2339327-A1.
 XX 15-MAR-2002.
 XX 19-MAR-2001; 2001CA-2339327.
 XX 15-SEP-2000; 2000US-0677374.
 XX (THOR/) THORNTON J C.
 XX (KAYW/) KAY W W.
 XX (BURI/) BURIAN J.
 XX (KUYZ/) KUYZIK M A.
 XX Thornton JC, Kay WW, Burian J, Kuzyk MA;
 XX WPI; 2002-455221/49.
 XX N-PSDB; ABE52401.
 XX Inducing immunity in fin fish to Rickettsial septicaemia, comprises
 PT administration of an outer surface lipoprotein (OspA) of a bacterial
 PT strain, as a vaccine -
 XX Claim 15; Fig 2; 55pp; English.
 XX The invention describes a method of protecting a poikilothermic fish
 CC against infection by the bacterial pathogen *Piscirickettsia salmonis*
 CC comprising administering either intraperitoneally, by immersion or
 CC orally, an immunogenic amount of principal antigen, the OspA (outer
 CC surface lipoprotein), its variants, non-lipidated form or antigenic
 CC peptides derived or synthesized with or without an adjuvant. The new
 CC method is used to provide an outer surface lipoprotein (OspA) of
 CC bacterial strain *Piscirickettsia salmonis* as a vaccine to induce immunity
 CC in fin-fish against Rickettsial septicaemia and other related
 CC Rickettsial diseases caused by either a virus, bacteria or parasite.
 CC This is the amino acid sequence of the *Piscirickettsia salmonis* outer
 CC surface lipoprotein, OspA, used in the creation of the vaccine described
 CC in the invention.
 XX Sequence 162 AA;
 SQ
 Query Match 61.0%; Score 815; DB 23; Length 162;
 Best Local Similarity 98.8%; Pred. No. 6.4e-72;
 Matches 158; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 97 RCLQSSLLIIISVFLVGCNQNFPSRQEVGAATGAVVGVAGQLFGKSGRVSMAIGGAVL 156
 DB 3 RCLQSSLLIIISVFLVGCNQNFPSRQEVGAATGAVVGVAGQLFGKSGRVSMAIGGAVL 62
 QY 157 GGLIGSKIGSQMDQDDKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQBR 216
 DB 63 GGLIGSKIGSQMDQDDKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQBR 122
 QY 217 RQCYCREFOQKAMIAQKQEIYGTACPDGRWQVISTEK 256
 DB 123 RQCYCREFOQKAMIAQKQEIYGTACPDGRWQVISTEK 162
 RESULT 8
 AAE05746
 ID AAE05746 standard; Protein; 190 AA.
 XX
 AC AAE05746;
 XX
 DT 24-SEP-2001 (first entry)
 XX
 DE Clostridium cellulovorans cellulose binding domain-180 (CBD-180).
 XX

KW Polysaccharide modification; polysaccharide binding domain; PBD; paper;
 KW yarn; fiber; textile; biological crosslinker; mechanical property;
 KW Clostridium cellulovorans cellulose binding domain; wet strength;
 KW durability; elasticity; CBDclos; cellulose binding protein A; CBP A;
 KW CBD-180.
 XX Clostridium cellulovorans.
 XX WO200134091-A2.
 XX 17-MAY-2001.
 XX 02-NOV-2000; 2000WO-IL00708.
 XX 08-NOV-1999; 99US-0164140.
 XX 18-NOV-1999; 99US-0166389.
 XX (CBDT-) CBD TECHNOLOGIES LTD.
 XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 XX Levy I, Nussinovitch A, Shoseyov O;
 XX WPI; 2001-457121/49.
 XX N-PSDB; AAD11043.
 XX Preparation of a polysaccharide containing material having at least one
 PT desired structural, chemical, physical, electrical and/or mechanical
 PT property -
 XX Example 1.2; Fig 1e-1g; 121pp; English.
 XX The present invention relates to methods and compositions for cross-
 CC linking and/or modifying the properties of polysaccharide materials.
 CC The method involves treating the polysaccharide structure with a
 CC polysaccharide binding domain (PBD) fusion protein. The method is
 CC used to alter the structural, chemical, physical, electrical and
 CC mechanical properties of polysaccharide materials such as paper,
 CC yarns, fibers and textiles, using biological crosslinking agents.
 CC The polysaccharide containing materials have improved mechanical
 CC properties such as wet strengths, durability and elasticity. The PBD
 CC reagent is applied in the forming stage in fusing paper manufacture
 CC which eliminates the sizing step. The use of a biological crosslinker
 CC improves the recyclability of paper products. The PBD reagent maintains
 CC the fine fibers in a slurry therefore resulting in better recovery of
 CC raw materials. The PBD molecules are eluted by strong alkaline conditions
 CC which enhances the ability of the alkaline glue used in binding to
 CC penetrate paper.
 CC The present sequence is Clostridium cellulovorans cellulose binding
 CC domain (CBD)-180. CBD-180 contains 180 amino acids of C. cellulovorans
 CC CBD (CBDclos) of cellulose binding protein A (CBP A).
 XX Sequence 190 AA;
 SQ
 Query Match 36.5%; Score 487; DB 22; Length 190;
 Best Local Similarity 46.8%; Pred. No. 1.4e-39;
 Matches 116; Conservative 12; Mismatches 30; Indels 90; Gaps 7;
 QY 1 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVVKVRYTSDGTQGTQTFWCDHAGAL 60
 DB 1 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVVKVRYTSDGTQGTQTFWCDHAGAL 60
 QY 61 LGNSYVDNTSKVTANFVKETASPTSTYDYLDPDSHMRGCLQGSSLLIIISVFLVGCNQNF 120
 DB 61 LGNSYVDNTSKVTANFVKETASPTSTYDYLDPDSHMRGCLQGSSLLIIISVFLVGCNQNF 92
 QY 121 ROEVGAATGAVVGVAGQLFGKSGRVSMAIGGAVLGLIGSKIGSQMDQDDKIKLNQSL 180
 DB 93 -----FGFASGRATL-----KGFQFIIQ----- 111
 QY 181 EKVKAGQVTR--WRNPDTGNSYSVEPVRTYQRYNKQRRQYCRFQOKAMTAGQKQEIY 238
 DB 112 -----GRITKSDWSNYTQNDYSFDASSSTPVNPF-----KVTGYIGGAK--VL 153

QY 239 GTACPPD 246
 Db 154 GTA-PGPD 160

RESULT 9
 AAE05749
 ID AAE05749 standard; Protein; 328 AA.
 AC AAE05749;
 XX
 XX 24-SEP-2001 (first entry)
 DE Chimeric S peptide-cellulose binding domain-S protein.
 KW Polysaccharide modification; polysaccharide binding domain; PBD; paper;
 KW yarn; fiber; textile; biological crosslinker; mechanical property;
 KW wet strength; durability; elasticity; cellulose binding domain; CBD;
 KW chimeric protein; S peptide-cellulose binding domain-S protein;
 KW Spet-CBD-Sprot; bovine.
 XX Chimeric - Clostridium cellulovorans.
 OS Chimeric - Bos sp.
 XX
 XX Key Location/Qualifiers
 FH Region 30..208
 FT /note= "This region is derived from C. cellulovorans"
 FT Region 226..326
 FT /note= "This region is derived from bovine"
 FT Misc-difference 327
 FT /label= Unknown
 FT /note= "Encoded by TAG"
 FT Misc-difference 327..328
 FT /note= "These residues are absent in the sequence shown
 as SEQ ID NO: 10 in the sequence listing"
 XX
 XX WO200134091-A2.
 XX 17-MAY-2001.
 XX
 XX 02-NOV-2000; 2000WO-IL00708.
 XX 08-NOV-1999; 99US-0164140.
 PR 18-NOV-1999; 99US-0166389.
 XX (CBT-) CBD TECHNOLOGIES LTD.
 PA (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 XX Levy I, Nussinovitch A, Shoseyov O;
 XX WPI: 2001-457121/49.
 DR N-PSDB; AAD11046.
 XX
 XX Preparation of a polysaccharide containing material having at least one
 PT desired structural, chemical, physical, electrical and/or mechanical
 PT property -
 XX Example 4; Fig 4b-4g; 121pp; English.
 XX
 CC The present invention relates to methods and compositions for cross-
 CC linking and/or modifying the properties of polysaccharide materials.
 CC The method involves treating the polysaccharide structure with a
 CC polysaccharide binding domain (PBD) fusion protein. The method is
 CC used to alter the structural, chemical, physical, electrical and
 CC mechanical properties of polysaccharide materials such as paper,
 CC yarns, fibers and textiles, using biological crosslinking agents.
 CC The polysaccharide containing materials have improved mechanical
 CC properties such as wet strengths, durability and elasticity. The PBD
 CC reagent is applied in the forming stage in fluting paper manufacture
 CC which eliminates the sizing step. The use of a biological crosslinker
 CC improves the recyclability of paper products. The PBD reagent maintains
 CC the fine fibers in a slurry therefore resulting in better recovery of
 CC raw materials. The PBD molecules are eluted by strong alkaline conditions

CC which enhances the ability of the alkaline glue used in binding to
 CC penetrate paper.
 CC The present sequence is S peptide-cellulose binding domain-S protein
 CC (Spet-CBD-Sprot), a fusion protein derived from Clostridium cellulovorans
 CC and bovine.
 CC Note: This sequence is stated to be the same as that shown as
 CC SEQ ID NO: 10 in the sequence listing of the specification. However
 CC this sequence has 2 additional residues at its C-terminal.
 XX
 XX SQ Sequence 328 AA;
 Query Match 36.5%; Score 487; DB 22; Length 328;
 Best Local Similarity 46.8%; Pred. No. 3.1e-39;
 Matches 116; Conservative 12; Mismatches 30; Indels 90; Gaps 7;
 QY 1 MSVEPYNKNKAQNTSITPIIKITNTSDNLNDVKRYVYTSQGTGOTFCWDHAGAL 60
 Db 35 MSVEFYNKNKAQNTSITPIIKITNTSDNLNDVKRYVYTSQGTGOTFCWDHAGAL 94
 QY 61 LGNSYVDNTSKVTANFVKETASPTSTYDTYLDPSHMRGCLGSSLIISVFLVGCAQNF 120
 Db 95 LGNSYVDNTSKVTANFVKETASPTSTYDTYVE----- 126
 QY 121 RQEVGAATGAVVGGVAGQLFGKSGRVSMAGGAVLGLGSLGKSGQSDQDKIKLNOSL 180
 Db 127 -----FGFASGRATL-----KKGQFITIQ----- 145
 QY 181 EKVAGQVTR--WRNPDTGNSVSEPVETVYQRYNKRQOYCRFFQKAMTAGOKQEIY 238
 Db 146 -----GRITKGDWSNYQTNDYSFDASSSTPVNP-----KVTGYIGGAK--VL 187
 QY 239 GTACPPD 246
 Db 188 GTA-PGPD 194

RESULT 10
 AAE05747
 ID AAE05747 standard; Protein; 341 AA.
 XX
 XX AC AAE05747;
 XX
 XX 24-SEP-2001 (first entry)
 XX Clostridium cellulovorans CBD cross linker protein (CCP).
 DE
 XX Polysaccharide modification; polysaccharide binding domain; PBD; paper;
 KW yarn; fiber; textile; biological crosslinker; mechanical property;
 KW wet strength; durability; elasticity; cellulose binding domain; CBD;
 KW CBD cross linker protein; CCP.
 XX
 XX OS Clostridium cellulovorans.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 340..341
 FT /note= "Encoded by CCATAGGAT"
 FT Misc-difference 341
 FT /note= "This residue is absent in the sequence shown
 as SEQ ID NO: 6 in the sequence listing"
 XX
 XX WO200134091-A2.
 XX 17-MAY-2001.
 XX
 XX 02-NOV-2000; 2000WO-IL00708.
 PR 08-NOV-1999; 99US-0164140.
 PR 18-NOV-1999; 99US-0166389.
 XX (CBT-) CBD TECHNOLOGIES LTD.
 PA (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
 XX Levy I, Nussinovitch A, Shoseyov O;

XX WPI; 2001-457121/49.
 DR N-PSDB; AAD11044.
 XX
 PT Preparation of a polysaccharide containing material having at least one
 PT desired structural, chemical, physical, electrical and/or mechanical
 PT property -
 XX
 XX
 PS Example 1.2; Fig 2b-2e; 121pp; English.
 XX
 CC The present invention relates to methods and compositions for cross-
 CC linking and/or modifying the properties of polysaccharide materials.
 CC The method involves treating the polysaccharide structure with a
 CC polysaccharide binding domain (PBD) fusion protein. The method is
 CC used to alter the structural, chemical, physical, electrical and
 CC mechanical properties of polysaccharide materials such as paper,
 CC yarns, fibers and textiles, using biological crosslinking agents.
 CC The polysaccharide containing materials have improved mechanical
 CC properties such as wet strengths, durability and elasticity. The PBD
 CC reagent is applied in the forming stage in fluting paper manufacture
 CC which eliminates the sizing step. The use of a biological crosslinker
 CC improves the recyclability of paper products. The PBD reagent maintains
 CC the fine fibers in a slurry therefore resulting in better recovery of
 CC raw materials. The PBD molecules are eluted by strong alkaline conditions
 CC which enhances the ability of the alkaline glue used in binding to
 CC penetrate paper.
 CC The present sequence is Clostridium cellulovorans cellulose binding
 CC domain (CBD) cross linker protein (CCP).
 CC Note: This sequence is stated to be the same as that shown as
 CC SEQ ID NO: 6 in the sequence listing of the specification. However
 CC this sequence has an additional residue at its C-terminal.
 XX
 SQ Sequence 341 AA;

Query Match 36.5%; Score 487; DB 22; Length 341;
 Best Local Similarity 46.8%; Pred. No. 3.3e-39;
 Matches 116; Conservative 12; Mismatches 39; Indels 90; Gaps 7;
 SQ
 QY 1 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVKVRYYYTSDGTQGTFCWDHAGAL 60
 DB 1 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVKVRYYYTSDGTQGTFCWDHAGAL 60
 QY 61 LGNSYVDNTSKVTANFVKETASPTSTYDTPSPHMRGCLQGSSLIISVFLVGCAQNFS 120
 DB 61 LGNSYVDNTSKVTANFVKETASPTSTYDTPSPHMRGCLQGSSLIISVFLVGCAQNFS 120
 QY 121 ROEVGAATGAVVGGVAGQLFGKGSRVSMAGGAVLGLIGSKIGQSMDDQDKIKLNQSL 180
 DB 121 ROEVGAATGAVVGGVAGQLFGKGSRVSMAGGAVLGLIGSKIGQSMDDQDKIKLNQSL 180
 QY 93 -----FGFASGRATL-----KKGQFITIQ----- 111
 DB 93 -----FGFASGRATL-----KKGQFITIQ----- 111
 QY 181 EKVKGAVQVTR--WRNPDTGNSYSVEPVRTYQRYNKRERQYCRFQOKAMTAGOKQEIY 238
 DB 112 -----GRITKSDWSNYTQNDYSFDASSSTPVVNP-----KVTGYTGGAK--VL 153
 QY 239 GTACQPD 246
 DB 154 GFA-FGPD 160

RESULT 11
 AAR63634
 ID AAR63634 standard; Protein; 162 AA.
 XX
 AC AAR63634;
 XX
 DT 25-MAR-2003 (updated)
 DT 05-JUN-1995 (first entry)
 XX
 DE Cellulose binding domain.
 XX Cellulose binding domain; CBD.
 XX Clostridium cellulovorans.

XX WO9424158-A1.
 XX
 PD 27-OCT-1994.
 XX
 PF 14-APR-1994; 94WO-US04132.
 XX
 PR 14-APR-1993; 93US-0048164.
 XX
 PA (REGC) UNIV CALIFORNIA.
 PA (YISS) YISSUM RES & DEV CO.
 XX
 PI Doi RH, Goldstein NA, Shoseyov O, Shpiegl I;
 DR WPI; 1994-341767/42.
 DR N-PSDB; AAO72917.
 XX
 PT Isolated cellulose binding domain and fusion proteins - with
 PT applications, including drug delivery, affinity separations, and
 PT diagnostic techniques
 XX
 PS Claim 1; Fig 1; 125pp; English.

CC AAO72917/Re3634 is a novel isolated cellulose binding domain. It pref.
 CC binds crystalline with a Kd ranging from 1.5 microm - 0.5 microm,
 CC pref. with a Kd of <1.2 microm, esp. <1.0 microm. The protein and
 CC the encoding DNA are claimed, as is a CBD fusion protein (FP)
 CC comprising the CBD and a second protein. The second protein is pref.
 CC Protein A, heparinase, a hormone or an enzyme capable of degrading
 CC an environmental pollutant, or an HSP, HSP antibody, cross-reactive
 CC HSP-related protein or an antigenic portion of this. The CBD and FP
 CC may be used in drug delivery, affinity sepsns, and diagnostic
 CC techniques. CBD nucleic acid may be obtd. from a variety of cell
 CC sources that produce CBDs that bind with high affinity and in a
 CC reversible manner or that produce CBD encoding mRNA. The preferred
 CC source of CBD encoding nucleic acid is C. cellulovorans.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX

SQ Sequence 162 AA;

Query Match 36.2%; Score 483; DB 15; Length 162;
 Best Local Similarity 51.7%; Pred. No. 2.8e-39;
 Matches 106; Conservative 9; Mismatches 14; Indels 76; Gaps 4;
 SQ
 QY 1 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVKVRYYYTSDGTQGTFCWDHAGAL 60
 DB 6 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVKVRYYYTSDGTQGTFCWDHAGAL 65
 QY 61 LGNSYVDNTSKVTANFVKETASPTSTYDTPSPHMRGCLQGSSLIISVFLVGCAQNFS 120
 DB 66 LGNSYVDNTSKVTANFVKETASPTSTYDTPSPHMRGCLQGSSLIISVFLVGCAQNFS 120
 QY 121 ROEVGAATGAVVGGVAGQLFGKGSRVSMAGGAVLGLIGSKIGQSMDDQDKIKLNQSL 180
 DB 98 -----FGFASGRATL-----KKGQFITIQ----- 116
 QY 181 EKVKGAVQVTR--WRNPDTGNSYSVE 203
 DB 117 -----GRITKSDWSNYTQNDYSFD 136

RESULT 12
 AAE05745
 ID AAE05745 standard; Protein; 163 AA.
 XX
 AC AAE05745;
 XX
 DT 24-SEP-2001 (first entry)
 XX
 DE Clostridium cellulovorans cellulose binding domain (CBDc1os).
 XX Polysaccharide modification; polysaccharide binding domain; PBD; paper;
 XX yarn; fiber; textile; biological crosslinker; mechanical property;
 KW

KW Clostridium cellulovorans cellulose binding domain; wet strength;
XX durability; elasticity; CBDelos; cellulose binding protein A; CBP A.
OS Clostridium cellulovorans.
XX WO200134091-A2.
XX
PD 17-MAY-2001.
XX
XX
XX
XX 02-NOV-2000; 2000WO-IL00708.
XX
XX 08-NOV-1999; 99US-0164140.
XX 18-NOV-1999; 99US-0166389.
XX
XX (CBDT-) CBD TECHNOLOGIES LTD.
XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX
XX Levy I, Nussinovitch A, Shoseyov O;
XX
XX WPI; 2001-457121/49.
XX N-PSDB; AAD11042.
XX
XX Preparation of a polysaccharide containing material having at least one
PT desired structural, chemical, physical, electrical and/or mechanical
PT property -
XX
XX Example 1.1; Page 111-112; 121pp; English.
XX
XX The present invention relates to methods and compositions for cross-
CC linking and/or modifying the properties of polysaccharide materials.
CC The method involves treating the polysaccharide structure with a
CC polysaccharide binding domain (PBD) fusion protein. The method is
CC used to alter the structural, chemical, physical, electrical and
CC mechanical properties of polysaccharide materials such as paper,
CC yarns, fibers and textiles, using biological crosslinking agents.
CC The polysaccharide containing materials have improved mechanical
CC properties such as wet strengths, durability and elasticity. The PBD
CC reagent is applied in the forming stage in fluting paper manufacture
CC which eliminates the sizing step. The use of a biological crosslinker
CC improves the recyclability of paper products. The PBD reagent maintains
CC the fine fibers in a slurry therefore resulting in better recovery of
CC raw materials. The PBD molecules are eluted by strong alkaline conditions
CC which enhances the ability of the alkaline glue used in binding to
CC penetrate paper.
CC The present sequence is Clostridium cellulovorans cellulose binding
CC domain (CBDelos) of cellulose binding protein A (CBP A).
XX
XX Sequence 163 AA;
SQ
Query Match 36.2%; Score 483; DB 22; Length 163;
Best Local Similarity 51.7%; Pred. No. 2.8e-39;
Matches 106; Conservative 9; Mismatches 14; Indels 76; Gaps 4;
QY 1 MSVEFNSKSAQTNSITPIIKITNTSDSLNDLVKRVRYTSDGTQQTWCDHAGAL 60
DB 7 MSVEFNSKSAQTNSITPIIKITNTSDSLNDLVKRVRYTSDGTQQTWCDHAGAL 66
QY 61 LGNSYVDNTSKVTANFVKETASPTSTYDTYLDPSHMRGCLGSSLLIISVFLVGCAQNF 120
DB 67 LGNSYVDNTSKVTANFVKETASPTSTYDIIVE----- 98
QY 121 ROEFGAATGAVGVAGQQLFGKSGRVSMAIGAVLGGILGSKIGQSMDDQDKIKLNSL 180
DB 99 -----FGFASGRATL-----KKGQFIQ----- 117
QY 181 EKVKAQOVTR--WRNPDTGNSVSVE 203
DB 118 -----GRITKSDWSNTYTQNDYSFD 137
RESULT 13
AAE05748
ID AAE05748 standard; Protein; 428 AA.

XX AAE05748;
XX
XX 24-SEP-2001 (first entry)
XX
XX Clostridium cellulovorans protein A-cellulose binding domain.
XX
XX Polysaccharide modification; polysaccharide binding domain; PBD; paper;
KW yarn; fiber; textile; biological crosslinker; mechanical property;
KW wet strength; durability; elasticity; protein A-cellulose binding domain;
KW ProtA-CBD.
XX
XX Clostridium cellulovorans.
XX
XX Key Location/Qualifiers
FT Region 1..263
FT /label= Protein A
FT /note= "This region is derived from pRIT2T vector"
FT Region 265..426
FT /label= CBP_A_protein
FT Misc-difference 427
FT /label= Unknown
FT /note= "Encoded by TAG"
FT Misc-difference 427..428
FT /note= "These residues are absent in the sequence shown
FT as SEQ ID NO: 8 in the sequence listing"
XX
XX WO200134091-A2.
XX
XX 17-MAY-2001.
XX
XX 02-NOV-2000; 2000WO-IL00708.
XX
XX 08-NOV-1999; 99US-0164140.
XX 18-NOV-1999; 99US-0166389.
XX
XX (CBDT-) CBD TECHNOLOGIES LTD.
XX (YISS) YISSUM RES DEV CO HEBREW UNIV JERUSALEM.
XX
XX Levy I, Nussinovitch A, Shoseyov O;
XX
XX WPI; 2001-457121/49.
XX N-PSDB; AAD11045.
XX
XX Preparation of a polysaccharide containing material having at least one
PT desired structural, chemical, physical, electrical and/or mechanical
PT property -
XX
XX Example 1.3; Fig 3b-3g; 121pp; English.
XX
XX The present invention relates to methods and compositions for cross-
CC linking and/or modifying the properties of polysaccharide materials.
CC The method involves treating the polysaccharide structure with a
CC polysaccharide binding domain (PBD) fusion protein. The method is
CC used to alter the structural, chemical, physical, electrical and
CC mechanical properties of polysaccharide materials such as paper,
CC yarns, fibers and textiles, using biological crosslinking agents.
CC The polysaccharide containing materials have improved mechanical
CC properties such as wet strengths, durability and elasticity. The PBD
CC reagent is applied in the forming stage in fluting paper manufacture
CC which eliminates the sizing step. The use of a biological crosslinker
CC improves the recyclability of paper products. The PBD reagent maintains
CC the fine fibers in a slurry therefore resulting in better recovery of
CC raw materials. The PBD molecules are eluted by strong alkaline conditions
CC which enhances the ability of the alkaline glue used in binding to
CC penetrate paper.
CC The present sequence is Clostridium cellulovorans Protein A-cellulose
CC binding domain (ProtA-CBD). This sequence contains a part of CBP A
CC protein and a part of protein A derived from pRIT2T vector sequence.
CC Note: This sequence is stated to be the same as that shown as
CC SEQ ID NO: 8 in the sequence listing of the specification. However
CC this sequence has 2 additional residues at its C-terminal.
XX

SQ Sequence 428 AA;
 Query Match 36.2%; Score 483; DB 22; Length 428;
 Best Local Similarity 51.7%; Pred. No. 1.1e-38;
 Matches 106; Conservative 9; Mismatches 14; Indels 76; Gaps 4;
 QY 1 MSVEFYNSKSAQTSITPIIKITNTSDSLNDLVKRYVYTSBGTQGTQFCDHAGAL 60
 DB 270 MSVEFYNSKSAQTSITPIIKITNTSDSLNDLVKRYVYTSBGTQGTQFCDHAGAL 329
 QY 61 LGNSYVDNTSKVTANFVKETASPTSTYDYLDPHMRGCLQSSLIISVFLVGCAQNPFS 120
 DB 330 LGNSYVDNTSKVTANFVKETASPTSTYDYLDPHMRGCLQSSLIISVFLVGCAQNPFS 361
 QY 121 ROEVAATGAVGVAGQVGFGRVSGRVSMAIGGAVLGLIGSGIGSQMDQDQKIKLNQSL 180
 DB 362 -----FGASGRATL-----KQGFITIQ----- 380
 QY 181 EKVKAQVTR--WRNPDTGNSYSVE 203
 DB 381 -----GRITKSDWSNYTQNDYSFD 400

RESULT 14
 AAY39952
 ID AAY39952 standard; Protein; 382 AA.
 AC AAY39952;
 XX
 DT 13-DEC-1999 (first entry)
 DE Gaussia luciferase fusion protein sequence.

KW Luciferase; green fluorescent protein; GFP; screening assay; diagnosis;
 KW bioluminescence-generating system; toy; cosmetic; fairy dust; beverage;
 KW body paint; squirt gun; balloon; slimy play material; soap; toothpaste;
 KW fusion protein.

XX Gaussia sp.
 XX WO9949019-A2.
 XX 30-SEP-1999.
 XX 26-MAR-1999; 99WO-US06698.
 XX 27-MAR-1998; 98US-0079624.
 PR 15-JUN-1998; 98US-0089367.
 PR 01-OCT-1998; 98US-0102939.
 XX (PROL-) PROLUME LTD.
 PA (BRYA/) BRYAN B J.
 XX Bryan BJ, Szent-Gyorgyi C;
 XX WPI; 1999-580443/49.
 DR N-PSDB; AA227550.
 XX New isolated Renilla mulleri, Gaussia and Pleuromamma luciferase and
 PT Renilla and Ptilosarcus green fluorescent protein nucleic acids -
 XX Disclosure; Page 222-223; 233pp; English.

XX This sequence represents a luciferase of the invention. The invention
 CC relates to Renilla mulleri, Gaussia and Pleuromamma luciferase and
 CC Renilla and Ptilosarcus green fluorescent protein (GFP) nucleic acids and
 CC proteins. The luciferases and GFPs can be used in
 CC bioluminescence-generating systems, assays, screening methods, diagnostic
 CC method and articles of manufacture. They can be expressed using
 CC e.g. Bacterial, yeast, fungal, plant, insect or animal cells. The Renilla
 CC mulleri, Gaussia and Pleuromamma luciferase or Renilla or Ptilosarcus
 CC GFPs can be used in e.g. toys, cosmetics, fountains, personal care items,
 CC fairy dust, beverages, soft drinks, foods, textile products, bubbles,

CC balloons, personal items, dentrifices, soaps, body paints, bubble bath,
 CC ink or paper products. In particular, they can be used in e.g. squirt
 CC guns, pellet guns, finger paints, foot bags, greeting cards, slimy play
 CC material, clothing, bubble making toys, bath powders, cosmetics, body
 CC lotions, gels, body powders, body creams, toothpastes, mouthwashes,
 CC soaps, body paints, bubble bath, inks, wrapping paper, gelatins, icings,
 CC frostings, greeting cards, beer, wine, champagne, soft drinks, ice cubes,
 CC ice, dry ice or fountains. The nucleic acids can also be used to produce
 CC transgenic fish and plants.

SQ Sequence 382 AA;

Query Match 36.0%; Score 480.5; DB 20; Length 382;
 Best Local Similarity 62.6%; Pred. No. 1.7e-38;
 Matches 107; Conservative 10; Mismatches 33; Indels 21; Gaps 4;

QY 1 MSVEFYNSKSAQTSITPIIKITNTSDSLNDLVKRYVYTSBGTQGTQFCDHAGAL 60
 DB 1 MSVEFYNSKSAQTSITPIIKITNTSDSLNDLVKRYVYTSBGTQGTQFCDHAGAL 60
 QY 61 LGNSYVDNTSKVTANFVKETASPTSTYDYLDPHMRGCLQSSLIISVFLVGCAQ 117
 DB 61 LGNSYVDNTSKVTANFVKETASPTSTYDYLDPHMRGCLQSSLIISVFLVGCAQ 120
 QY 118 NFSRQ-----EVGAATGAVGVAGQVGFGRVSGRVSMAIGGA-VLGGIGS 162
 DB 121 NVTQNDYSFDASSSTPVNPKVTGY-----IGGAKVLGTAPGS 159

RESULT 15
 AAE13383
 ID AAE13383 standard; Protein; 382 AA.
 XX
 AC AAE13383;
 DT 12-FEB-2002 (first entry)
 DE Gaussia species CBD-luciferase fusion protein.

XX Green fluorescent protein; GFP; bioluminescence generating system; toy;
 KW luciferase; finger paint; slimy play material; fishing lure; sparkler;
 KW doll; balloon; personal care item; cosmetic; bath powder; body cream;
 KW tooth paste; mouth wash; soap; body paint; ornamental transgenic plant;
 KW bubble bath; fountain; fairy dust; food; gelatin; icing; frosting; beer;
 KW wine; champagne; milk; soft drink; ice cube; ice; foot bag; beverage;
 KW dry ice; textile; clothing; paper product; greeting card; wrapping paper;
 KW fusion protein.

XX Gaussia sp.
 XX WO200168824-A2.
 XX 20-SEP-2001.
 XX 15-MAR-2001; 2001WO-US08277.
 XX 15-MAR-2000; 2000US-189691P.
 XX (PROL-) PROLUME LTD.
 PA (BRYA/) BRYAN B J.

XX Bryan BJ, Szent-Gyorgyi C, Szczepaniak W;
 XX WPI; 2002-010561/01.
 DR N-PSDB; AAD22201.
 XX Nucleic acids encoding Renilla reniformis green fluorescent proteins,
 PT useful in diagnostic bioluminescence procedures -
 XX Disclosure; Page 163-164; 175pp; English.
 PS The patent discloses sea pansy (Renilla reniformis) green fluorescent
 CC proteins (GFP) and their corresponding polynucleotides. The invention

CC also relates to sequences of the bioluminescence generating system
CC (e.g. luciferase). R. reniformis GFP are used in diagnostic methods
CC and in the production of novelty items such as toys (e.g. squirt gun,
CC pellet gun, toy "Halloween" eggs, toy cigarettes, board/card game
CC toys), finger paints, slimy play material, bubbles in bubble making
CC toys, fishing lures, dolls, sparklers, magic wand toys, balloons,
CC personal care item (e.g. cosmetic, bath powders, body creams, tooth
CC pastes, mouth wash, soaps, body paints, bubble bath), ornamental
CC transgenic plants, fountain, fairy dust, food (gelatins, icings,
CC frosting, beers, wine, champagne, milk, soft drinks, ice cubes, ice,
CC dry ice, beverage), textile (foot bag, clothing) and/or paper product
CC (greeting cards, wrapping paper). The present sequence is CBD-luciferase
CC fusion protein from *Gaussia* species.
XX
SQ

Sequence 382 AA;

Query Match 36.0%; Score 480.5; DB 23; Length 382;
Best Local Similarity 62.8%; Pred. No. 1.7e-38;
Matches 107; Conservative 10; Mismatches 33; Indels 21; Gaps 4;
QY 1 MSVEFYNSKSAQTNSITPIIKITNTSDSLNDVKVRYTSDGTQGTWCDFHAGAL 60
DB 1 MSVEFYNSKSAQTNSITPIIKITNTSDSLNDVKVRYTSDGTQGTWCDFHAGAL 60
QY 61 LGNSYVDNTSKVTANFVKETASPTSTYDYLDPFMRGCL---QGSLLIIISVFLVGCAQ 117
DB 61 LGNSYVDNTSKVTANFVKETASPTSTYDYLDPFMRGCL---QGSLLIIISVFLVGCAQ 117
QY 118 NPSRQ-----EVGAATCAVVGVGAGQLFGKSGRVSMAIGGA-VLGGLIGS 162
DB 121 NYQTNDYSGDASSSTFVVPKVTGY-----IGGAKVLGTAFGS 159

Search completed: November 5, 2003, 20:12:51
Job time : 47.4249 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 20:11:02 ; Search time 16.3592 Seconds
(without alignments)
662.108 Million cell updates/sec

Title: US-09-677-374-6

Perfect score: 1335

Sequence: 1 MSVEFYNSKSAQTNSITPI.....IYTACPOQDGRMQVISTEK 256

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/prodata/1/aaa/5A-COMB.pep:*
2: /cgn2_6/prodata/1/aaa/5B-COMB.pep:*
3: /cgn2_6/prodata/1/aaa/6A-COMB.pep:*
4: /cgn2_6/prodata/1/aaa/6B-COMB.pep:*
5: /cgn2_6/prodata/1/aaa/PCTUS-COMB.pep:*
6: /cgn2_6/prodata/1/aaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	480.5	36.0	382	US-09-277-716-22	Sequence 22, Appl
2	480.5	36.0	382	US-09-609-161B-22	Sequence 22, Appl
3	477	35.7	162	US-08-048-164A-2	Sequence 2, Appl
4	477	35.7	162	US-08-460-462-2	Sequence 2, Appl
5	477	35.7	162	US-08-460-457-2	Sequence 2, Appl
6	477	35.7	162	US-08-460-458-2	Sequence 2, Appl
7	477	35.7	162	US-08-460-455-2	Sequence 2, Appl
8	477	35.7	162	US-08-330-394A-2	Sequence 2, Appl
9	477	35.7	163	US-09-006-636-7	Sequence 7, Appl
10	477	35.7	163	US-09-006-632-7	Sequence 7, Appl
11	477	35.7	163	US-09-325-274-7	Sequence 7, Appl
12	472	35.4	154	US-08-330-394A-29	Sequence 29, Appl
13	472	35.4	156	US-08-330-394A-22	Sequence 22, Appl
14	245	18.4	167	PCT-US95-13813-9	Sequence 9, Appl
15	245	18.4	476	US-09-339-159B-4	Sequence 4, Appl
16	245	18.4	493	US-09-198-956-10	Sequence 10, Appl
17	245	18.4	493	US-09-198-955A-12	Sequence 12, Appl
18	245	18.4	493	US-09-694-531-12	Sequence 12, Appl
19	245	18.4	493	US-09-670-141-10	Sequence 10, Appl
20	149.5	11.2	1426	US-09-136-574A-43	Sequence 43, Appl
21	148.5	11.1	616	US-09-136-574A-47	Sequence 47, Appl
22	148.5	11.1	1751	US-09-136-574A-44	Sequence 44, Appl
23	127.5	9.6	192	US-09-252-991A-22817	Sequence 22817, A
24	122.5	9.2	551	US-09-033-537A-1	Sequence 1, Appl
25	111.5	8.4	165	US-09-252-991A-19701	Sequence 19701, A
26	100.5	7.5	387	US-09-252-991A-23334	Sequence 23334, A
27	100.5	7.5	700	US-07-862-580B-2	Sequence 2, Appl

28	99.5	7.5	306	4	US-09-252-991A-21754	Sequence 21754, A
29	98	7.3	922	4	US-09-198-452A-15	Sequence 15, Appl
30	96.5	7.2	112	4	US-09-252-991A-31960	Sequence 31960, A
31	94.5	7.1	217	4	US-09-328-352-7068	Sequence 7068, Ap
32	92.5	6.9	199	4	US-09-328-352-7802	Sequence 7802, Ap
33	92.5	6.9	1034	4	US-09-252-991A-26658	Sequence 26658, A
34	91.5	6.9	1415	4	US-09-252-991A-26438	Sequence 26438, A
35	90	6.7	126	4	US-09-328-352-8009	Sequence 8009, Ap
36	88.5	6.6	431	4	US-09-107-532A-4816	Sequence 4816, Ap
37	87	6.5	790	4	US-09-328-352-4283	Sequence 4283, Ap
38	86.5	6.5	127	4	US-09-252-991A-28397	Sequence 28397, A
39	86.5	6.5	730	4	US-09-328-352-4442	Sequence 4442, Ap
40	85.5	6.4	257	4	US-10-092-246-2	Sequence 2, Appl
41	85	6.4	258	4	US-09-328-352-4253	Sequence 4253, Ap
42	85	6.4	513	4	US-09-134-001C-4469	Sequence 4469, Ap
43	85	6.4	1021	1	US-07-910-760-12	Sequence 12, Appl
44	85	6.4	1021	1	US-08-440-519-12	Sequence 12, Appl
45	85	6.4	1021	4	US-08-440-549-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1

US-09-277-716-22

; Sequence 22, Application US/092777716A

; Patent No. 6232107

; GENERAL INFORMATION:

; APPLICANT: Bryan, Bruce,

; APPLICANT: Szent-Gyorgyi, Christopher

; APPLICANT: PROMUNE, LTD.

; TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE

; CURRENT APPLICATION NUMBER: US/09/277,716A

; CURRENT FILING DATE: 1999-03-26

; EARLIER APPLICATION NUMBER: 60/102,939

; EARLIER FILING DATE: 1998-10-01

; EARLIER APPLICATION NUMBER: 60/089,367

; EARLIER FILING DATE: 1998-06-15

; EARLIER APPLICATION NUMBER: 60/079,624

; EARLIER FILING DATE: 1998-03-27

; NUMBER OF SEQ ID NOS: 32

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 22

; LENGTH: 382

; TYPE: PRT

; ORGANISM: Artificial Sequence: fusion protein

; FEATURE:

; OTHER INFORMATION: Cellulose Binding Domain (CBD)-Gaussia luciferase fusion protein

US-09-277-716-22

Query Match 36.0%; Score 480.5; DB 3; Length 382;

Best Local Similarity 62.6%; Pred. No. 5.8e-41;

Matches 107; Conservative 10; Mismatches 33; Indels 21; Gaps 4;

QY 1 MSVEFYNSKSAQTNSITPIIKITNTSDNLNDLVKRYYYTSDGTQGTFCWDHAGAL 60

Db 1 MSVEFYNSKSAQTNSITPIIKITNTSDNLNDLVKRYYYTSDGTQGTFCWDHAGAL 60

QY 61 LGNSYVDNTSKVTANFVKETASPTSTYDYLDPDSMRGCL-----QGSSLIISVFLVGCQAQ 117

Db 61 LGNSYVDNTSKVTANFVKETASPTSTYDYLDPDSMRGCL-----QGSSLIISVFLVGCQAQ 117

QY 118 NPSRQ-----EVGAATGAVGVGAGLFGKSGRVSMAIGGA-VLGGGLGS 162

Db 121 NYTQTDNSYFDASSSTPVVNPVKVTGY-----IGGAKVLGTAPGS 159

RESULT 2

US-09-609-161B-22

; Sequence 22, Application US/09609161B

; Patent No. 6436682

; GENERAL INFORMATION:

; APPLICANT: Bryan, Bruce

APPLICANT: Szent-Gyorgyi, Christopher
APPLICANT: PROLUME, LTD.
TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE LUC
TITLE OF INVENTION: AND FLUORESCENT PROTEINS AND THE USE THEREOF IN DIAGNOSTICS, HIG
TITLE OF INVENTION: SCREENING AND NOVELTY ITEMS
FILE REFERENCE: 24729-121B
CURRENT APPLICATION NUMBER: US/09/609,161B
CURRENT FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: 09/277,716
PRIOR FILING DATE: 1999-03-26
PRIOR APPLICATION NUMBER: 60/102,939
PRIOR FILING DATE: 1998-10-01
PRIOR APPLICATION NUMBER: 60/089,367
PRIOR FILING DATE: 1998-06-15
PRIOR APPLICATION NUMBER: 60/079,624
PRIOR FILING DATE: 1998-03-27
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 22
LENGTH: 382
TYPE: PRT
ORGANISM: Artificial Sequence: fusion protein
FEATURE:
OTHER INFORMATION: Cellulose Binding Domain (CBD)-Gaussia luciferase fusion protein
US-09-609-161B-22

Query Match 36.0%; Score 480.5; DB 4; Length 382;
Best Local Similarity 62.8%; Pred. No. 5.8e-41;
Matches 107; Conservative 10; Mismatches 33; Indels 21; Gaps 4;
QY 1 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVVKVRYTSDGTQGTFCWCDHAGAL 60
DB 1 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVVKVRYTSDGTQGTFCWCDHAGAL 60
QY 61 LGNSYVDNTSKVTANFVKETASPTSTYDYLDPHMRGCL---QGSLLIIISVFLVGCQA 117
DB 61 LGNSYVDNTSKVTANFVKETASPTSTYDYLDPHMRGCL---QGSLLIIISVFLVGCQA 117
QY 118 NFRSQ-----EVCANTCAVVGAGQLFGKGRVSNMAGGA-VLGLIGS 162
DB 121 NYQTNDYSPDASSSTPVNPKVTGY-----IGGAKVLGTAPGS 159

RESULT 3
US-08-048-164A-2
Sequence 2, Application US/08048164A
Patent No. 5496934
GENERAL INFORMATION:
APPLICANT: Shoseyov, Oded
APPLICANT: Shpiegl, Itai
APPLICANT: Goldstein, Marc A.
APPLICANT: Doi, Roy H.
TITLE OF INVENTION: CELLULOSE BINDING DOMAIN
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESS: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/048,164A
FILING DATE: 14-APR-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mirock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7809-006
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 162 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-460-462-2

Query Match 35.7%; Score 477; DB 1; Length 162;

REFERENCE/DOCKET NUMBER: 7809-003
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 162 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-048-164A-2
Query Match 35.7%; Score 477; DB 1; Length 162;
Best Local Similarity 97.8%; Pred. No. 3.8e-41;
Matches 90; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVVKVRYTSDGTQGTFCWCDHAGAL 60
DB 6 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVVKVRYTSDGTQGTFCWCDHAGAL 65
QY 61 LGNSYVDNTSKVTANFVKETASPTSTYDYLDPHMRGCL---QGSLLIIISVFLVGCQA 117
DB 66 LGNSYVDNTSKVTANFVKETASPTSTYDYLDPHMRGCL---QGSLLIIISVFLVGCQA 117

RESULT 4
US-08-460-462-2
Sequence 2, Application US/08460462
Patent No. 5670623
GENERAL INFORMATION:
APPLICANT: Shoseyov, Oded
APPLICANT: Shpiegl, Itai
APPLICANT: Goldstein, Marc A.
APPLICANT: Doi, Roy H.
TITLE OF INVENTION: METHODS OF USE OF CELLULOSE BINDING DOMAIN PROTEINS
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESS: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,462
FILING DATE: concurrently herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/048,164
FILING DATE: 14-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Mirock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 7809-006
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 162 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-460-462-2

Query Match 35.7%; Score 477; DB 1; Length 162;

Best Local Similarity 97.8%; Pred. No. 3.8e-41; Mismatches 0; Indels 0; Gaps 0;
Matches 90; Conservative 2;

QY 1 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVVKVRYVYTSQGTQGTWCDHAGAL 60
Db 6 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVVKVRYVYTSQGTQGTWCDHAGAL 65

QY 61 LGNSYVDNTSKVTANFVKETASPTSTYDTYLD 92
Db 66 LGNSYVDNTSKVTANFVKETASPTSTYDTYVE 97

RESULT 5

US-08-460-457-2
; Sequence 2, Application US/08460457
; Patent No. 5719044
; GENERAL INFORMATION:
; APPLICANT: Shoseyov, Oded
; APPLICANT: Shpiegl, Itai
; APPLICANT: Goldstein, Marc A.
; APPLICANT: Doi, Roy H.
; TITLE OF INVENTION: CELLULOSE BINDING DOMAIN FUSION PROTEINS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: concurrently herewith
; APPLICATION NUMBER: US/08/460,457
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/048,164
; FILING DATE: 14-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7809-008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 869-8864/9741
; TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 162 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-460-457-2

Query Match 35.7%; Score 477; DB 1; Length 162;
Best Local Similarity 97.8%; Pred. No. 3.8e-41;
Matches 90; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVVKVRYVYTSQGTQGTWCDHAGAL 60
Db 6 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVVKVRYVYTSQGTQGTWCDHAGAL 65

QY 61 LGNSYVDNTSKVTANFVKETASPTSTYDTYLD 92
Db 66 LGNSYVDNTSKVTANFVKETASPTSTYDTYVE 97

RESULT 6

US-08-460-458-2

; Sequence 2, Application US/08460458
; Patent No. 5738984

; GENERAL INFORMATION:
; APPLICANT: Shoseyov, Oded
; TITLE OF INVENTION: KITS AND METHODS OF DETECTION USING CELLULOSE BINDING DOMAIN
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: concurrently herewith
; APPLICATION NUMBER: US/08/460,458
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/048,164
; FILING DATE: 14-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7809-007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 869-8864/9741
; TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 162 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-460-458-2

Query Match 35.7%; Score 477; DB 1; Length 162;
Best Local Similarity 97.8%; Pred. No. 3.8e-41;
Matches 90; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVVKVRYVYTSQGTQGTWCDHAGAL 60
Db 6 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVVKVRYVYTSQGTQGTWCDHAGAL 65

QY 61 LGNSYVDNTSKVTANFVKETASPTSTYDTYLD 92
Db 66 LGNSYVDNTSKVTANFVKETASPTSTYDTYVE 97

RESULT 7

US-08-460-455-2
; Sequence 2, Application US/08460455
; Patent No. 5837814

; GENERAL INFORMATION:
; APPLICANT: Shoseyov, Oded
; APPLICANT: Shpiegl, Itai
; APPLICANT: Goldstein, Marc A.
; APPLICANT: Doi, Roy H.
; TITLE OF INVENTION: CELLULOSE BINDING DOMAIN PROTEINS
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,455
; FILING DATE: concurrently herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/048,164
; FILING DATE: 14-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7809-009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 162 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-460-455-2

Query Match 35.7%; Score 477; DB 2; Length 162;
Best Local Similarity 97.8%; Pred. No. 3.8e-41;
Matches 90; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVKVRYVYTSQTGQTFCWDHAGAL 60
Db 6 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVKVRYVYTSQTGQTFCWDHAGAL 65

QY 61 LGNSYVDNTSKVTANFVKETASPTSTYDYLD 92
Db 66 LGNSYVDNTSKVTANFVKETASPTSTYDYVE 97

RESULT 8
US-08-330-394A-2
; Sequence 2, Application US/08330394A
; Patent No. 5856201
; GENERAL INFORMATION:
; APPLICANT: Shoseyov, Oded
; APPLICANT: Yosef, Karney
; APPLICANT: Shpiegl, Itai
; APPLICANT: Goldstein, Marc A.
; APPLICANT: Doi, Roy H.
; TITLE OF INVENTION: METHODS OF DETECTION USING THE
; TITLE OF INVENTION: CELLULOSE BINDING DOMAIN
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 AVENUE OF THE AMERICAS
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/330,394A
; FILING DATE: 27-OCT-1994
; PRIOR APPLICATION DATA: CIP OF PCT/US94/04132
; APPLICATION NUMBER:
; FILING DATE: 14-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/460,455
; FILING DATE: concurrently herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/048,164
; FILING DATE: 14-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7809-009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 162 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-460-455-2

Query Match 35.7%; Score 477; DB 2; Length 162;
Best Local Similarity 97.8%; Pred. No. 3.8e-41;
Matches 90; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVKVRYVYTSQTGQTFCWDHAGAL 60
Db 6 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVKVRYVYTSQTGQTFCWDHAGAL 65

QY 61 LGNSYVDNTSKVTANFVKETASPTSTYDYLD 92
Db 66 LGNSYVDNTSKVTANFVKETASPTSTYDYVE 97

RESULT 9
US-09-006-636-7
; Sequence 7, Application US/09006636
; Patent No. 6005092
; GENERAL INFORMATION:
; APPLICANT: Shoseyov, Oded
; APPLICANT: Shani, Ziv
; TITLE OF INVENTION: ARABIDOPSIS THALIANA ENDO-1,4-BETA-
; TITLE OF INVENTION: GLUCANASE GENE, PROMOTER AND PROTEIN
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/006,636
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Baldwin, Geraldine F.
; REGISTRATION NUMBER: 31,232
; REFERENCE/DOCKET NUMBER: 7809-019
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 163 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-006-636-7

Query Match 35.7%; Score 477; DB 3; Length 163;
Best Local Similarity 97.8%; Pred. No. 3.8e-41;
Matches 90; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVKVRYYYTSDGTQQTFCWDHAGAL 60
 Db 7 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVKVRYYYTSDGTQQTFCWDHAGAL 66
 QY 61 LGNSYVDNTSKVTANFVKETASPTSTYDYLD 92
 Db 67 LGNSYVDNTSKVTANFVKETASPTSTYDYVE 98

RESULT 10
 US-09-006-632-7
 ; Sequence 7, Application US/09006632
 ; Patent No. 6184440
 ; GENERAL INFORMATION:
 ; APPLICANT: Shoseyov, Oded
 ; APPLICANT: Shani, Ziv
 ; APPLICANT: Shigel, Etai
 ; TITLE OF INVENTION: TRANSGENIC PLANTS OF ALTERED
 ; TITLE OF INVENTION: MORPHOLOGY
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds LLP
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSEQ Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/006,632
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Baldwin, Geraldine F.
 ; REGISTRATION NUMBER: 31,232
 ; REFERENCE/DOCKET NUMBER: 7809-018
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 790-9090
 ; TELEFAX: (212) 869-8864
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 163 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-09-006-632-7

Query Match 35.7%; Score 477; DB 3; Length 163;
 Best Local Similarity 97.8%; Pred. No. 3.8e-41;
 Matches 90; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVKVRYYYTSDGTQQTFCWDHAGAL 60
 Db 7 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVKVRYYYTSDGTQQTFCWDHAGAL 66
 QY 61 LGNSYVDNTSKVTANFVKETASPTSTYDYLD 92
 Db 67 LGNSYVDNTSKVTANFVKETASPTSTYDYVE 98

RESULT 11
 US-09-325-274-7
 ; Sequence 7, Application US/09325274
 ; Patent No. 6323023
 ; GENERAL INFORMATION:
 ; APPLICANT: Shoseyov, Oded
 ; APPLICANT: Shani, Ziv

; TITLE OF INVENTION: ARABIDOPSIS THALIANA ENDO-1,4-BETA-
 ; TITLE OF INVENTION: GLUCANASE GENE, PROMOTER AND PROTEIN
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds LLP
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: NY
 ; COUNTRY: USA
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSEQ Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/325,274
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 09/006,636
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Baldwin, Geraldine F.
 ; REGISTRATION NUMBER: 31,232
 ; REFERENCE/DOCKET NUMBER: 7809-019
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 790-9090
 ; TELEFAX: (212) 869-8864
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 163 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-09-325-274-7

Query Match 35.7%; Score 477; DB 4; Length 163;
 Best Local Similarity 97.8%; Pred. No. 3.8e-41;
 Matches 90; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVKVRYYYTSDGTQQTFCWDHAGAL 60
 Db 7 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVKVRYYYTSDGTQQTFCWDHAGAL 66
 QY 61 LGNSYVDNTSKVTANFVKETASPTSTYDYLD 92
 Db 67 LGNSYVDNTSKVTANFVKETASPTSTYDYVE 98

RESULT 12
 US-08-330-394A-29
 ; Sequence 29, Application US/08330394A
 ; Patent No. 5856201
 ; GENERAL INFORMATION:
 ; APPLICANT: Shoseyov, Oded
 ; APPLICANT: Yosef, Karmey
 ; APPLICANT: Shpiegel, Itai
 ; APPLICANT: Goldstein, Marc A.
 ; APPLICANT: Doi, Roy H.
 ; TITLE OF INVENTION: METHODS OF DETECTION USING THE
 ; TITLE OF INVENTION: CELLULOSE BINDING DOMAIN
 ; NUMBER OF SEQUENCES: 33
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: PENNIE & EDMONDS
 ; STREET: 1155 AVENUE OF THE AMERICAS
 ; CITY: NEW YORK
 ; STATE: NEW YORK
 ; COUNTRY: U.S.A.
 ; ZIP: 10036
 ; COMPUTER READABLE FORM:

```
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/330,394A
/ FILING DATE: 27-OCT-1994
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: CIP OF PCT/US94/04132
/ FILING DATE: 14-APR-1994
/ ATTORNEY/AGENT INFORMATION:
/ NAME: MISROCK, S. LESLIE
/ REFERENCE/DOCKET NUMBER: 7809-005
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 790-9090
/ TELEX: 66441 PENNIE
/ INFORMATION FOR SEQ ID NO: 29:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 154 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-330-394A-29

Query Match      35.4%; Score 472; DB 2; Length 154;
Best Local Similarity 97.8%; Pred. No. 1.1e-40;
Matches 89; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 SVFYNKSAQNTSIPIIKITNTSDSLNLDVKNVRYTSDGTQGTFCWCHAGALL 61
Db 1 SVFYNKSAQNTSIPIIKITNTSDSLNLDVKNVRYTSDGTQGTFCWCHAGALL 60

QY 62 GNSYVDNTSKVTANFVKETASPTSTYDTYLD 92
Db 61 GNSYVDNTSKVTANFVKETASPTSTYDTYVE 91

/ US-08-330-394A-29

RESULT 13
US-08-330-394A-22
; Sequence 22, Application US/08330394A
; Patent No. 5856201
; GENERAL INFORMATION:
; APPLICANT: Shoseyov, Oded
; APPLICANT: Yosef, Karmey
; APPLICANT: Shpiegl, Itai
; APPLICANT: Goldstein, Marc A.
; APPLICANT: Doi, Roy H.
; TITLE OF INVENTION: METHODS OF DETECTION USING THE
; TITLE OF INVENTION: CELLULOSE BINDING DOMAIN
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 AVENUE OF THE AMERICAS
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/330,394A
; FILING DATE: 27-OCT-1994
; PRIOR APPLICATION DATA: CIP OF PCT/US94/04132
; APPLICATION NUMBER:
; FILING DATE: 14-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REFERENCE/DOCKET NUMBER: 7809-005
; TELECOMMUNICATION INFORMATION:
```

```
/ TELEPHONE: (212) 790-9090
/ TELEFAX: (212) 869-8864
/ TELEX: 66441 PENNIE
/ INFORMATION FOR SEQ ID NO: 22:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 156 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-330-394A-22

Query Match      35.4%; Score 472; DB 2; Length 156;
Best Local Similarity 97.8%; Pred. No. 1.2e-40;
Matches 89; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 SVFYNKSAQNTSIPIIKITNTSDSLNLDVKNVRYTSDGTQGTFCWCHAGALL 61
Db 1 SVFYNKSAQNTSIPIIKITNTSDSLNLDVKNVRYTSDGTQGTFCWCHAGALL 60

QY 62 GNSYVDNTSKVTANFVKETASPTSTYDTYLD 92
Db 61 GNSYVDNTSKVTANFVKETASPTSTYDTYVE 91

/ US-08-330-394A-22

RESULT 14
PCT-US95-13813-9
; Sequence 9, Application PC/TUS9513813
; GENERAL INFORMATION:
; APPLICANT: Yeda Research and Development Co. Ltd.
; APPLICANT: Ramot University Authority for Applied
; APPLICANT: Research and Industrial Development Ltd.
; APPLICANT: Technion Research and Development Foundation Ltd.
; APPLICANT: Bayer, Edward A.
; APPLICANT: Morag, Ely
; APPLICANT: Wilchek, Meir
; APPLICANT: Lamed, Raphael
; APPLICANT: Shoham, Yuval
; TITLE OF INVENTION: MODIFIED CELLULOSE-BINDING DOMAIN (CBD)
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street N.W., Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13813
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Browdy, Roger L.
; REGISTRATION NUMBER: 25,618
; REFERENCE/DOCKET NUMBER: BAYER=3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 167 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US95-13813-9
```

Query Match	18.4%;	Score 245;	DB 5;	Length 167;
Best Local Similarity	51.5%;	Pred. No. 2.4e-17;		
Matches 53;	Conservative 16;	Mismatches 30;	Indels 4;	Gaps 3;

Qy 1 MSVEFYNSKSAQTNSITPIKIITNTSDSLNLDNVKRYYYTSDGTQGQTFWCDHAGAL 60
: ||||| | |||| | ||| | : : : |||| : ||||| :
db 9 LKVEFYNSPDDTNSINPQKVTNTGSSAIDLKSLTRYYYTVDGQKDOTFWCDHA-AI 67

RESULT 15

```

US-09-339-159B-4
; Sequence 4, Application US/09339159B
; Patent No. 6566114
; GENERAL INFORMATION:
; APPLICANT: Kauppinen, Markus
; APPLICANT: Schulein, Martin
; APPLICANT: Schnorr, Kirk
; APPLICANT: Andersen, Lene
; APPLICANT: Bjørnvad, Mads
; TITLE OF INVENTION: No. 6566114el Mannanases
; FILE REFERENCE: 5440-204-US
; CURRENT APPLICATION NUMBER: US/09/339,159B
; CURRENT FILING DATE: 1999-06-24
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Bacillus
US-09-339-159B-4

```

Query Match 18.4%; Score 245; DB 4; Length 476;
Best Local Similarity 51.5%; Pred. No. 1.1e-16;
Matches 53; Conservative 16; Mismatches 30; Indels 4; Gaps 3;

Qy	1	MSVEFYSNKS	SA	TN	SI	TP	II	KI	KT	NT	SD	SL	ND	NV	KV	RY	YT	SD	TG	QT	GF	WC	DH	AG	AL	60
		:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	
Db	318	LKVEFYSN	SP	DT	NS	IN	PF	KV	TN	TG	SS	AI	DL	SK	LT	LR	Y	YT	VD	GK	QD	TF	WC	DH	-AI	376

Search completed: November 5, 2003, 20:16:47
Job time : 17.3592 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 20:15:12 ; Search time 30.0656 Seconds
(without alignments)
1462.395 Million cell updates/sec

Title: US-09-677-374-6

Perfect score: 1335

Sequence: 1 MSVEFYNSKSAQNTSITPI.....IYGTACPDQGRWQVISTEK 256

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 644079 seqs, 171749292 residues

Total number of hits satisfying chosen parameters: 644079

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PTC_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US05_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PTCUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1335	100.0	256	12	US-10-261-446-6
2	832	62.3	161	12	US-10-261-446-4
3	815	61.0	162	12	US-10-261-446-2
4	480.5	36.0	382	11	US-09-808-898-22
5	245	18.4	493	14	US-10-072-152-12
6	245	18.4	599	11	US-09-955-555A-29
7	127.5	9.6	182	15	US-10-127-032-169
8	127.5	9.6	1352	11	US-09-784-554B-2
9	127	9.5	741	15	US-10-156-761-8100
10	126.5	9.5	1350	11	US-09-784-554B-4
11	123.5	9.3	1043	11	US-09-917-384-6
12	123.5	9.3	1043	11	US-09-917-384-6
13	121	9.1	150	11	US-09-917-383-5
14	121	9.1	150	11	US-09-917-383-5
15	121	9.1	1228	11	US-09-917-384-1

16	121	9.1	1228	11	US-09-917-383-1	Sequence 1, Appli
17	120.5	9.0	154	11	US-09-917-378-4	Sequence 4, Appli
18	120.5	9.0	762	11	US-09-917-378-1	Sequence 1, Appli
19	115	8.6	957	15	US-10-155-400-1	Sequence 1, Appli
20	114	8.5	88	15	US-10-155-400-5	Sequence 5, Appli
21	114	8.5	89	15	US-10-155-400-4	Sequence 4, Appli
22	113	8.5	1621	15	US-10-185-990-10	Sequence 10, Appl
23	112	8.4	20	12	US-10-261-446-16	Sequence 16, Appl
24	98	7.3	922	9	US-09-886-468-19	Sequence 19, Appl
25	92	6.9	285	10	US-09-978-295A-28	Sequence 28, Appl
26	92	6.9	285	10	US-09-978-697-28	Sequence 28, Appl
27	92	6.9	285	10	US-09-978-192A-28	Sequence 28, Appl
28	92	6.9	285	10	US-09-999-832A-28	Sequence 28, Appl
29	92	6.9	285	11	US-09-978-189-28	Sequence 28, Appl
30	92	6.9	285	11	US-09-978-608A-28	Sequence 28, Appl
31	92	6.9	285	11	US-09-978-585A-28	Sequence 28, Appl
32	92	6.9	285	11	US-09-978-191A-28	Sequence 28, Appl
33	92	6.9	285	11	US-09-978-403A-28	Sequence 28, Appl
34	92	6.9	285	11	US-09-978-564A-28	Sequence 28, Appl
35	92	6.9	285	11	US-09-999-833A-28	Sequence 28, Appl
36	92	6.9	285	11	US-09-981-915A-28	Sequence 28, Appl
37	92	6.9	285	11	US-09-978-824-28	Sequence 28, Appl
38	92	6.9	285	11	US-09-918-585A-28	Sequence 28, Appl
39	92	6.9	285	11	US-09-978-423A-28	Sequence 28, Appl
40	92	6.9	285	11	US-09-978-193A-28	Sequence 28, Appl
41	92	6.9	285	11	US-09-999-830A-28	Sequence 28, Appl
42	92	6.9	285	11	US-09-978-757A-28	Sequence 28, Appl
43	92	6.9	285	11	US-09-978-187B-28	Sequence 28, Appl
44	92	6.9	285	11	US-09-978-643A-28	Sequence 28, Appl
45	92	6.9	285	12	US-09-978-375A-28	Sequence 28, Appl

ALIGNMENTS

RESULT 1

US-10-261-446-6
; Sequence 6, Application US/10261446
; Publication No. US20030165526A1
; GENERAL INFORMATION:
; APPLICANT: Kuzik, Michael A.
; APPLICANT: Burian, Jan
; APPLICANT: Kay, William W.
; APPLICANT: Thornton, Julian C.
; TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY IN FISH AGAINST
; TITLE OF INVENTION: RICKETTSIAL DISEASES, AND ASSOCIATED PREVENTATIVE THERAPY
; FILE REFERENCE: 4616-84094
; CURRENT APPLICATION NUMBER: US/10/261,446
; CURRENT FILING DATE: 2002-09-30
; PRIOR FILING DATE: US 09/677,374
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/154,437
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 256
; TYPE: PRP
; ORGANISM: *Piscirickettsia salmonis*
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (-95)..(-1)
US-10-261-446-6

Query Match 100.0%; Score 1335; DB 12; Length 256;
Best Local Similarity 100.0%; Pred. No. 1.4e-128;
Matches 256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSVEFYNSKSAQNTSITPIIKITNTSDNLNDVKRYVYTSDTGOTGQFWCDHAGAL 60
DB 1 MSVEFYNSKSAQNTSITPIIKITNTSDNLNDVKRYVYTSDTGOTGQFWCDHAGAL 60
QY 61 LGNSYVDNTSKVTANFVKETASPTSTYTLDPSPHMRGCLQSSLLIIISVFLVGCAQNF 120

Db 61 LGNSYDNTSKVTANFVKETASPTSTYDYLDPSHMRGCLQGSSLLIISVFLVGCAQNF 120
QY 121 ROEVAATGAVVGGVAGQLFGKSGRVSMAIGAVLGGIGSKIGSMDQDDKIKLNQSL 180
Db 121 ROEVAATGAVVGGVAGQLFGKSGRVSMAIGAVLGGIGSKIGSMDQDDKIKLNQSL 180
QY 181 EKVKAGVTRWRNPDTGNSYSVEPVRTYORYNKQERRQYCREFOOKAMIAQKOEIYGT 240
Db 181 EKVKAGVTRWRNPDTGNSYSVEPVRTYORYNKQERRQYCREFOOKAMIAQKOEIYGT 240
QY 241 ACPQDGRWQVISTEK 256
Db 241 ACPQDGRWQVISTEK 256

RESULT 2
US-10-261-446-4
; Sequence 4, Application US/10261446
; Publication No. US20030165526A1
; GENERAL INFORMATION:
; APPLICANT: Kuzik, Michael A.
; APPLICANT: Burian, Jan
; APPLICANT: Kay, William W.
; APPLICANT: Thornton, Julian C.
; TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY IN FISH AGAINST
; TITLE OF INVENTION: RICKETTSIAL DISEASES, AND ASSOCIATED PREVENTATIVE THERAPY
; FILE REFERENCE: 4616-64094
; CURRENT APPLICATION NUMBER: US/10/261,446
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US 09/677,374
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/154,437
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Piscirickettsia salmonis
US-10-261-446-4

Query Match 62.3%; Score 832; DB 12; Length 161;
Best Local Similarity 100.0%; Pred. No. 2.7e-77;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 96 MRGCLQGSSLLIISVFLVGCAQNFQSRQEVGAATGAVVGGVAGQLFGKSGRVSMAIGAV 155
Db 1 MRGCLQGSSLLIISVFLVGCAQNFQSRQEVGAATGAVVGGVAGQLFGKSGRVSMAIGAV 60
QY 156 LGGLGSKIGSMDQDDKIKLNQSLKVKAGVTRWRNPDTGNSYSVEPVRTYORYNKOE 215
Db 61 LGGLGSKIGSMDQDDKIKLNQSLKVKAGVTRWRNPDTGNSYSVEPVRTYORYNKOE 120
QY 216 RQOYCREFOOKAMIAQKOEIYGTACPDGRWQVISTEK 256
Db 121 RQOYCREFOOKAMIAQKOEIYGTACPDGRWQVISTEK 161

RESULT 3
US-10-261-446-2
; Sequence 2, Application US/10261446
; Publication No. US20030165526A1
; GENERAL INFORMATION:
; APPLICANT: Kuzik, Michael A.
; APPLICANT: Burian, Jan
; APPLICANT: Kay, William W.
; APPLICANT: Thornton, Julian C.
; TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY IN FISH AGAINST
; TITLE OF INVENTION: RICKETTSIAL DISEASES, AND ASSOCIATED PREVENTATIVE THERAPY
; FILE REFERENCE: 4616-64094
; CURRENT APPLICATION NUMBER: US/10/261,446
; CURRENT FILING DATE: 2002-09-30

; PRIOR APPLICATION NUMBER: US 09/677,374
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/154,437
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Piscirickettsia salmonis
US-10-261-446-2

Query Match 61.0%; Score 815; DB 12; Length 162;
Best Local Similarity 98.8%; Pred. No. 1.5e-75;
Matches 158; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 97 RGCLQGSSLLIISVFLVGCAQNFQSRQEVGAATGAVVGGVAGQLFGKSGRVSMAIGAVL 156
Db 3 RGCLQGSSLLIISVFLVGCAQNFQSRQEVGAATGAVVGGVAGQLFGKSGRVSMAIGAVL 62
QY 157 GGLGSKIGSMDQDDKIKLNQSLKVKAGVTRWRNPDTGNSYSVEPVRTYORYNKOE 216
Db 63 GGLGSKIGSMDQDDKIKLNQSLKVKAGVTRWRNPDTGNSYSVEPVRTYORYNKOE 122
QY 217 RQOYCREFOOKAMIAQKOEIYGTACPDGRWQVISTEK 256
Db 123 RQOYCREFOOKAMIAQKOEIYGTACPDGRWQVISTEK 162

RESULT 4
US-09-808-898-22
; Sequence 22, Application US/09808898
; Publication No. US20030092098A1
; GENERAL INFORMATION:
; APPLICANT: Bryan, Bruce
; APPLICANT: Szent-Gyorgyi, Christopher
; APPLICANT: Szent-Gyorgyi, William
; TITLE OF INVENTION: RENILLA RENIFORMIS FLOURESCENT PROTEINS, NUCLEIC ACIDS ENCODING 1
; TITLE OF INVENTION: FLOURESCENT PROTEINS AND THE USE THEREOF IN DIAGNOSTICS, HIGH TH
; TITLE OF INVENTION: SCREENING AND NOVELTY ITEMS
; FILE REFERENCE: 24729-128
; CURRENT APPLICATION NUMBER: US/09/808,898
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/189,691
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: 09/277,716
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 08/757,046
; PRIOR FILING DATE: 1996-11-25
; PRIOR APPLICATION NUMBER: 08/597,274
; PRIOR FILING DATE: 1996-02-06
; PRIOR APPLICATION NUMBER: 08/908,909
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 08/990,103
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Gaussia
US-09-808-898-22

Query Match 36.0%; Score 480.5; DB 11; Length 382;
Best Local Similarity 62.6%; Pred. No. 1.1e-40;
Matches 107; Conservative 10; Mismatches 33; Indels 21; Gaps 4;
QY 1 MSVEFYNNKSAQTNSITPILKITNTSDNLNDVKRYYYTSDGTGQTFWCDHAGAL 60
Db 1 MSVEFYNNKSAQTNSITPILKITNTSDNLNDVKRYYYTSDGTGQTFWCDHAGAL 60
QY 61 LGNSYDNTSKVTANFVKETASPTSTYDYLDPSHMRGCL---QGSSLLIISVFLVGCAQ 117

Db 61 LGNSYDNTSKVTANFVKETASPTSTYDTYVEFGFASGAATLKKQGFITIQGRITKSDWS 120
 QY 118 NFSRQ-----EVGAATGAVGVAGQLFGKSGRVSMAIGGA-VLGGLIGS 162
 Db 121 NYQTNDISFDASSSTPVVNPKVTGY-----IGGAKVLGTAPGS 159

RESULT 5
 US-10-072-152-12
 ; Sequence 12, Application US/10072152
 ; Publication No. US20020142438A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Andersen, Lene N.
 ; APPLICANT: Schulein, Martin
 ; APPLICANT: Lange, Niels E.
 ; APPLICANT: Bjornvad, Mads E.
 ; APPLICANT: Moller, Soren
 ; APPLICANT: Glad, Sanne O. S.
 ; APPLICANT: Kauppinen, Markus S.
 ; APPLICANT: Schnorr, Kirk
 ; APPLICANT: Kongsbak, Lars
 ; TITLE OF INVENTION: No. US20020142438A1el Pectate Lyases

; FILE REFERENCE: 5378.200-US
 ; CURRENT APPLICATION NUMBER: US/10/072,152
 ; CURRENT FILING DATE: 2002-02-07
 ; PRIOR APPLICATION NUMBER: US/09/198,955
 ; PRIOR FILING DATE: 1998-11-24
 ; PRIOR APPLICATION NUMBER: 1343/97
 ; PRIOR FILING DATE: 1997-11-24
 ; PRIOR APPLICATION NUMBER: 1344/97
 ; PRIOR FILING DATE: 1997-11-24
 ; PRIOR APPLICATION NUMBER: 60/067,249
 ; PRIOR FILING DATE: 1997-12-02
 ; PRIOR APPLICATION NUMBER: 60/067,240
 ; PRIOR FILING DATE: 1997-12-02
 ; PRIOR APPLICATION NUMBER: 09/073,684
 ; PRIOR FILING DATE: 1998-05-06
 ; PRIOR APPLICATION NUMBER: 09/184,217
 ; PRIOR FILING DATE: 1998-11-02
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 12
 ; LENGTH: 493
 ; TYPE: PRT
 ; ORGANISM: Clostridium thermocellum
 US-10-072-152-12

Query Match 18.4%; Score 245; DB 14; Length 493;
 Best Local Similarity 51.5%; Pred. No. 2.2e-16;
 Matches 53; Conservative 16; Mismatches 30; Indels 4; Gaps 3;
 QY 1 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVKKVYYVTS DGTGQTFCWCDHAGAL 60
 Db 335 LKVEFYNSPDSITNSINPQFKVNTGSSAIDLKSLTLRYYYTV DGGKQDFWCDHA-AI 393
 QY 61 LGN--SYVDNTSKVTANFVKETASPTSTYDTYLDPSHMRGCLQ 101
 Db 394 IGSNGSYNGITSNVKGTFVRKSSS-TNNADTYLEISFTGTLT 435

RESULT 6
 US-09-955-555A-29
 ; Sequence 29, Application US/09955555A
 ; Publication No. US20030027298A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bott, Richard R.
 ; APPLICANT: Clarkson, Kathleen A.
 ; APPLICANT: Fowler, Timothy
 ; APPLICANT: Liu, Chung-Cheng
 ; APPLICANT: Ward, Michael
 ; APPLICANT: Xia, Hai-Ying
 ; TITLE OF INVENTION: Enzymatic Array and Process of Making
 ; TITLE OF INVENTION: Same

; FILE REFERENCE: GC278-C3
 ; CURRENT APPLICATION NUMBER: US/09/955.555A
 ; CURRENT FILING DATE: 2001-09-17
 ; PRIOR APPLICATION NUMBER: US 08/559,958
 ; PRIOR FILING DATE: 1995-11-17
 ; PRIOR APPLICATION NUMBER: US 60/005,701
 ; PRIOR FILING DATE: 1995-10-17
 ; NUMBER OF SEQ ID NOS: 29
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 29
 ; LENGTH: 599
 ; TYPE: PRT
 ; ORGANISM: Clostridium thermocellum
 US-09-955-555A-29

Query Match 18.4%; Score 245; DB 11; Length 599;
 Best Local Similarity 51.5%; Pred. No. 2.9e-16;
 Matches 53; Conservative 16; Mismatches 30; Indels 4; Gaps 3;
 QY 1 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVKKVYYVTS DGTGQTFCWCDHAGAL 60
 Db 384 LKVEFYNSPDSITNSINPQFKVNTGSSAIDLKSLTLRYYYTV DGGKQDFWCDHA-AI 442
 QY 61 LGN--SYVDNTSKVTANFVKETASPTSTYDTYLDPSHMRGCLQ 101
 Db 443 IGSNGSYNGITSNVKGTFVRKSSS-TNNADTYLEISFTGTLT 484

RESULT 7
 US-10-127-032-169
 ; Sequence 169, Application US/10127032
 ; Publication No. US20030113742A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Whiteley, Marvin
 ; APPLICANT: Banger, M. Gita
 ; APPLICANT: Lory, Stephen
 ; APPLICANT: Greenberg, Everett Peter
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF
 ; TITLE OF INVENTION: BIOFILM FORMATION
 ; FILE REFERENCE: UI2-070CP
 ; CURRENT APPLICATION NUMBER: US/10/127,032
 ; CURRENT FILING DATE: 2002-04-19
 ; PRIOR APPLICATION NUMBER: US 60/285,190
 ; PRIOR FILING DATE: 2001-04-20
 ; PRIOR APPLICATION NUMBER: US 60/344,142
 ; PRIOR FILING DATE: 2001-10-24
 ; NUMBER OF SEQ ID NOS: 170
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 169
 ; LENGTH: 182
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-10-127-032-169

Query Match 9.6%; Score 127.5; DB 15; Length 182;
 Best Local Similarity 34.8%; Pred. No. 6e-05;
 Matches 32; Conservative 13; Mismatches 28; Indels 19; Gaps 2;
 QY 122 QEVGAATGAVGVAGQLFGKSGRVSMAIGGAVLGLIGSKIGQSMDDQDKI----- 174
 Db 70 QIAGTAIGAVVGGLLGNQIGGTTGKTIATVAGAVGGYAGNKKVQSGMOERDYYTTTETRC 129
 QY 175 -KLNQSLKLV-----KAGOVTRWRNP 194
 Db 130 STVHDSSEKVGVDYVKYMLDGKAGQIRMERDP 161

RESULT 8
 US-09-784-554B-2
 ; Sequence 2, Application US/09784554B
 ; Publication No. US20030032162A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Schnorr, Kirk

; APPLICANT: Jorgensen, Per Lina
; APPLICANT: Schulein, Martin
; TITLE OF INVENTION: FAMILY 44 XYLOGUCANASES
; FILE REFERENCE: 10017.200-US
; CURRENT APPLICATION NUMBER: US/09/784,554B
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1352
; TYPE: PRT
; ORGANISM: Paenibacillus polymyxa
US-09-784-554B-2

Query Match 9.6%; Score 127.5; DB 11; Length 1352;
Best Local Similarity 30.4%; Pred. No. 0.001;
Matches 28; Conservative 19; Mismatches 36; Indels 9; Gaps 2;
QY 3 VEFYNSKSAQTSITPIIKITNTSDSLNLDVVKVRYYYTSDGTQGTFCWCDHAGALLG 62
Db 1203 LQYRTADTKVNDNHLNPFQIVNKGTTSPINELKIRYYTIDGDRQTFNCDYATL--- 1259
QY 63 NSYVNTSKVTANFVKETASPTSTYTYLDPS 94
Db 1260 ----SCSKLNGKLVKMDKAATGA-DYYLEVS 1285

RESULT 9
US-10-156-761-8100
; Sequence 8100, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIEBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 8100
; LENGTH: 741
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-8100

Query Match 9.5%; Score 127; DB 15; Length 741;
Best Local Similarity 32.7%; Pred. No. 0.0005;
Matches 33; Conservative 15; Mismatches 43; Indels 10; Gaps 3;
QY 1 MSVEFYNSKSAQTSITPIIKITNTSDSLNLDVVKVRYYYTSD-GTQGTFCWCDHAGA 59
Db 594 LKVLKYNDSATDNNAIRPGLRIVNTGSGLDLSKVYRYTSDSGSPTVNAWCDAV 653
QY 60 LLGNSYVDNTSKVTANFVKETASPTSTYTYLDPSHMRGCL 100
Db 654 -----GCSNLSKVKVPLT-TPVPGADAYLEVGTGTL 685

RESULT 10
US-09-784-554B-4
; Sequence 4, Application US/09784554B
; Publication No. US20030032162A1
; GENERAL INFORMATION:
; APPLICANT: Schnorr, Kirk

; APPLICANT: Jorgensen, Per Lina
; APPLICANT: Schulein, Martin
; TITLE OF INVENTION: FAMILY 44 XYLOGUCANASES
; FILE REFERENCE: 10017.200-US
; CURRENT APPLICATION NUMBER: US/09/784,554B
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1350
; TYPE: PRT
; ORGANISM: Paenibacillus polymyxa
US-09-784-554B-4

Query Match 9.5%; Score 126.5; DB 11; Length 1350;
Best Local Similarity 31.5%; Pred. No. 0.0013;
Matches 29; Conservative 19; Mismatches 35; Indels 9; Gaps 2;
QY 3 VEFYNSKSAQTSITPIIKITNTSDSLNLDVVKVRYYYTSDGTQGTFCWCDHAGALLG 62
Db 1202 LQYRTADTKVNDNHLNPFQILNKGTTISVPINELKIRYYTIDGDRQTFNCDYA----- 1256
QY 63 NSYVNTSKVTANFVKETASPTSTYTYLDPS 94
Db 1257 ---VLSCSKLNGKLVKMDKAATGA-DYYLEVS 1284

RESULT 11
US-09-917-384-6
; Sequence 6, Application US/09917384
; Publication No. US20030096342A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: DECKER, STEPHEN R.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT CELLULOSE FROM ACIDOTHERMUS
; FILE REFERENCE: 40170.6US01
; CURRENT APPLICATION NUMBER: US/09/917,384
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1043
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Segment of
; OTHER INFORMATION: Guxa
US-09-917-384-6

Query Match 9.3%; Score 123.5; DB 11; Length 1043;
Best Local Similarity 24.6%; Pred. No. 0.0019;
Matches 58; Conservative 28; Mismatches 79; Indels 71; Gaps 9;
QY 1 MSVEFYNSKSAQTSITPIIKITNTSDSLNLDVVKVRYYYTSDGTQGTFCWCDHAGA 59
Db 481 LKQYKYNDSAPGDNQIKPGLQVLNTGSSVDLSVTYRYWYTRDGGSSSTLVYNCDAAM 540
QY 60 LLGNSYVDNTSKVTANFVKETASP-TSTYTYLDPSHMRGCLQ--GSS----- 104
Db 541 CCGN-----IRASP--GSVNPAFTADYILQLSFTGGTLAGGSGTGTQNRVKNKSDW 590
QY 105 -----LIISVFLVGCANFSPROEVGAATGAVVGGVAGQLFG 141
Db 591 SNFTETNDYSYGTNTTTFQDWTKVTVYVNGVLVWGTEDEVTPPSV--PTGLVVTGVS----- 644
QY 142 KSGRVSMAIGAVLGGILGSKIGQMDQDKIKLNQSLKVKAGQVTRWRNPDTG 197
Db 645 ---SSVSLAWNA-----STDNVGAHVNYVRNGVLVGQPTVTSFTDTG 684

```
RESULT 12
US-09-917-383-6
; Sequence 6, Application US/09917383
; Publication No. US20030104522A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: DECKER, STEPHEN R.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT CELLULOSE FROM ACIDOTHERMUS
; FILE REFERENCE: 40170.6US01
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1043
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Segment of
; OTHER INFORMATION: Guxa
US-09-917-383-6

Query Match          9.3%; Score 123.5; DB 11; Length 1043;
Best Local Similarity 24.6%; Pred. No. 0.0019;
Matches 58; Conservative 28; Mismatches 79; Indels 71; Gaps 9;

QY      1 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVKVRYVYTSQTGQTFW-CDHAGA 59
      481 LKVQYKNDSAPGDNQIKPGLQLVNTGSSVDLSVTYVYFTRDGGSSTLVYNCDAWAM 540
      60 LLGNSYVDNTSKVTANFVKETASP-TSTYDYLDPESHMRGCLQ--GSS----- 104
      541 GCGN-----IRASFPATPTADTYLQLSFTGTLAAGSGSTGEIQNRVKNKSDW 590
      105 -----LIIISVFLGCAQNSRQEVGAATGAVVGVAGQLFG 141
      591 SNFTETNYSYNTNTTFQDWTKVYVGVNGLVMTGEDVTTPSV--PTGLVWTGVS----- 644
      142 KSGSRVSNMAGVNLGGLIGSKIGOSMDQDDKILNQSLKVKAGQVTRWRNPDTG 197
      645 ----SSVSLAWNA-----STDNVGVAHVYVNGVLVQGPVTWTSFTDITG 684

RESULT 13
US-09-917-384-5
; Sequence 5, Application US/09917384
; Publication No. US20030096342A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: DECKER, STEPHEN R.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT CELLULOSE FROM ACIDOTHERMUS
; FILE REFERENCE: 40170.6US01
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Segment of
; OTHER INFORMATION: Guxa
US-09-917-384-5
```

```
Query Match          9.1%; Score 121; DB 11; Length 150;
Best Local Similarity 32.4%; Pred. No. 0.00021;
Matches 33; Conservative 19; Mismatches 38; Indels 12; Gaps 4;

QY      1 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVKVRYVYTSQTGQTFW-CDHAGA 59
      5 LKVQYKNDSAPGDNQIKPGLQLVNTGSSVDLSVTYVYFTRDGGSSTLVYNCDAWAM 64
      60 LLGNSYVDNTSKVTANFVKETASP-TSTYDYLDPESHMRGCL 100
      65 GCGN-----IRASFPATPTADTYLQLSFTGTL 96

RESULT 14
US-09-917-383-5
; Sequence 5, Application US/09917383
; Publication No. US20030104522A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: DECKER, STEPHEN R.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT CELLULOSE FROM ACIDOTHERMUS
; FILE REFERENCE: 40170.6US01
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Segment of
; OTHER INFORMATION: Guxa
US-09-917-383-5

Query Match          9.1%; Score 121; DB 11; Length 150;
Best Local Similarity 32.4%; Pred. No. 0.00021;
Matches 33; Conservative 19; Mismatches 38; Indels 12; Gaps 4;

QY      1 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVKVRYVYTSQTGQTFW-CDHAGA 59
      5 LKVQYKNDSAPGDNQIKPGLQLVNTGSSVDLSVTYVYFTRDGGSSTLVYNCDAWAM 64
      60 LLGNSYVDNTSKVTANFVKETASP-TSTYDYLDPESHMRGCL 100
      65 GCGN-----IRASFPATPTADTYLQLSFTGTL 96

RESULT 15
US-09-917-384-1
; Sequence 1, Application US/09917384
; Publication No. US20030096342A1
; GENERAL INFORMATION:
; APPLICANT: DING, SHI-YOU
; APPLICANT: ADNEY, WILLIAM S.
; APPLICANT: VINZANT, TODD B.
; APPLICANT: DECKER, STEPHEN R.
; APPLICANT: HIMMEL, MICHAEL E.
; TITLE OF INVENTION: THERMAL TOLERANT CELLULOSE FROM ACIDOTHERMUS
; FILE REFERENCE: 40170.6US01
; CURRENT FILING DATE: 2001-07-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1228
; TYPE: PRT
```

```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Segment of
; OTHER INFORMATION: Guxa
US-09-917-384-1

Query Match      9.1%; Score 121; DB 11; Length 1228;
Best Local Similarity 32.4%; Pred. NO. 0.0042;
Matches 33; Conservative 19; Mismatches 38; Indels 12; Gaps 4;

QY      1 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVKKVYYTSDGTQGTW-CDHAGA 59
Db      588 LKQYKNNDSAPGDNQIKPGLQVNTGSSVDLSTVTVWFTRDGGSTLVINCDAAM 647

QY      60 LLGNSYVDNTSKVTANFVKETASP-TSTYDYLDPSHMRGCL 100
Db      648 GCGN-----IRASP--GSVNPATPTADTYLQLSFTGGTL 679
```

Search completed: November 5, 2003, 20:27:20
Job time : 31.0656 secs

Db 120 -RNSTGOYCREYTOTVIGGKOOKAYGNACRQPDGOWQVYN 159

C33971

A;Reference number: S25767
A;Accession: S25767
A;Molecule type: DNA
A;Residues: 1216-1232, 'Y', 1234-1241, 'T', 1243-1319, 'R', 1321-1615, 1617-1854 <F03>
A;Cross-references: EMBL:X67406
R;Seguin, P.
submitted to the EMBL Data Library, August 1992
A;Reference number: S28659
A;Accession: S28659

A:Molecule type: DNA
A:Residues: 1216-1232,'Y',1234-1241,'T',1243-1319,'R',1321-1615,1617-1768,'R',1770-1854
A:Cross-references: EMBL:X67406
R:Fujino, T.; Beguin, P.; Aubert, J.P.
J. Bacteriol. 175, 1891-1899, 1993
A:Title: Organization of a Clostridium thermocellum gene cluster encoding the cellulose
e.
A:Reference number: Z18847; MUID:93209931; PMID:8458832
A:Accession: T18261
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1821-1854 <PU2>
A:Cross-references: EMBL:X67506; NID:g296879; PID:g296880; PIDN:CAA47840.1
C:Genetics:
A:Gene: cipa

Query Match 18.4%; Score 245; DB 2; Length 1854;
Best Local Similarity 51.5%; Pred. No. 7.5e-12;
Matches 53; Conservative 16; Mismatches 30; Indels 4; Gaps 3;
QY 1 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVVKVRYVYTS DGTGQTFCWDHAGAL 60
Db 369 LKVEFNSPSTNSINQFKVTNGSSAIDLKLTLYYYTV DQKQDTFCWDHA-AI 427
QY 61 LGN--SYVDNTSKVTANFVKETASPTSTYDYLDP SHMRGCLQ 101
Db 428 IGSNGSYNGITSNVKTGFVKMSS--INNADTYLEISFTG GTLE 469

RESULT 10
T30433
scaffolding protein - Clostridium josui
C:Species: Clostridium josui
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 21-Jul-2000
C:Accession: T30433
R:Kakuuchi, M.; Isui, A.; Suzuki, K.; Fujino, T.; Fujino, E.; Kimura, T.; Karita, S.; Sa
J. Bacteriol. 180, 4303-4308, 1998
A:Title: Cloning and DNA sequencing of the genes encoding Clostridium josui scaffolding
e.
A:Reference number: Z20837; MUID:98361925; PMID:9696784
A:Accession: T30433
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1162 <KAK>
A:Cross-references: EMBL:AB004845; NID:g3445476; PIDN:BAA32429.1; PID:g3445477
C:Genetics:
A:Gene: cipa

Query Match 17.5%; Score 233.5; DB 2; Length 1162;
Best Local Similarity 34.0%; Pred. No. 3.8e-11;
Matches 66; Conservative 31; Mismatches 74; Indels 23; Gaps 7;
QY 1 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVVKVRYVYTS DGTGQTFCWDHAGAL 60
Db 33 ISVQFNNGSSPTSSSSSYARFVKVTNTSGSPINLADLKLRYFTQDENKQMTFCWDHAGYL 92
QY 61 LGNSYVDNTSKVTANFVKETASPTSTYDYLDP--SHMRGCL--QGSSLLIISVELVGCAQ 117
Db 93 SGNVYMDVTSKVGTF-NEVSPAVTNADHYLEVALSSDAGSLPAGSIEIQIRFARNDS 151
QY 118 NFSRQEVGAATGAVVGGVAGOLFGRGS-----GRVSMAGVAVLGGIGSKIQSQMDQDDK 173
Db 152 NFDQSDNWSYTSA-----GSYMDWQKIAAFVGGTLYVG--STFNGDDNPTQDP 197
QY 174 IKLNQSLKVKVKAQ 187
Db 198 -KISPTSISAKQ 210

RESULT 11
C97012
probably cellulosomal scaffolding protein precursor, secreted, cellulose-binding and cof
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C:Accession: C97012
R:Nolling, J.; Breton, G.; Onelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cl
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: C97012
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1483 <KUR>
A:Cross-references: GB:AB001437; PIDN:AAK78886.1; PID:gl5023809; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC0910

Query Match 16.7%; Score 223.5; DB 2; Length 1483;
Best Local Similarity 28.4%; Pred. No. 3.4e-10;
Matches 57; Conservative 22; Mismatches 49; Indels 73; Gaps 4;
QY 1 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVVKVRYVYTS DGTGQTFCWDHAGAL 60
Db 37 VOIQPADTSTTWTMTIAPKFINNTGAPLDLTTLTKURYFTADGTQDENFWCDHAGWL 96
QY 61 LGNSYVDNTSKVTANFVKETASPTSTYDYLDP SHMRGCLQGSSLLIISVELVGCAQNF 120
Db 97 NGYNYQTITSNVVGTFV-AMDNATATADHYLEISFNG----- 133
QY 121 RQEVGAATGAVVGGVAGOLFGRGSRVSMAGVAVLGGIGSKIQSQMDQDDKIKLNQSL 180
Db 134 -----AGQL-----DAGSSLEVCQKRVAKND-- 153
QY 181 EKVAGQVTRWRNPDTGNSYS 201
Db 154 -----WSNYDQSDNYS 164

RESULT 12
A47704
endoglucanase I (EC 3.2.1.-) Cell - Clostridium thermocellum
C:Species: Clostridium thermocellum
C:Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 15-Oct-1999
C:Accession: A47704
R:Hazlewood, G.P.; Davidson, K.; Laurie, J.I.; Huskisson, N.S.; Gilbert, H.J.
J. Gen. Microbiol. 139, 307-316, 1993
A:Title: Gene sequence and properties of CellI, a family E endoglucanase from Clostridium
A:Reference number: A47704; MUID:93171873; PMID:8436949
A:Accession: A47704
A>Status: preliminary
A:Molecule type: DNA; protein
A:Residues: 1-879 <HAG>
A:Cross-references: GB:L04735; NID:g144807; PIDN:AAA20892.1; PID:g144808
A:Note: sequence extracted from NCBI backbone (NCBIN:125637, NCBI:P:125638)
C:Keywords: glycosidase; hydrolase

Query Match 14.1%; Score 188.5; DB 2; Length 879;
Best Local Similarity 28.4%; Pred. No. 1.4e-07;
Matches 59; Conservative 20; Mismatches 48; Indels 81; Gaps 9;
QY 3 VEFYNSKSAQTNSITPIIKITNTSDSLNLDVVKVRYVYTS DGTGQTFCWDHAGALLG 62
Db 742 LQYANGAGATGNSINPRFKIINNGTKAINLSVKIRYYTKEGGASQNFCDWSSA--G 799
QY 63 NSYVDNTSKVTANFVKETASPTSTYDYLDP SHMRGCLQGSSLLIISVELVGCAQNF 122
Db 800 N-----SNVTGNFF-NLSSPKRGAT-----CL----- 821
QY 123 EVGAATGAVVGGVAGOLFGRGSRVSMAGVAVLGGIGSKIQSQMDQDDKIKLNQSL 182
Db 822 EVG-----FGGAG--TLDPGGSV-----EVOIRFSKED 848
QY 183 VKAGQVTRWRNPDTGNSYSVEPVTYQR 210

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 20:06:36 ; Search time 10.1693 Seconds
(without alignments)
1183.846 Million cell updates/sec

Title: US-09-677-374-6

Perfect score: 1335

Sequence: 1 MSVEFYNSKSAQTNSITPI.....IYGTACPDGRWQVISTEK 256

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	487	36.5	1848	1 CBPA_CLOCL	P38058 clostridium
2	288	21.6	159	1 17KD_RICPR	P16624 rickettsia
3	286	21.4	159	1 17KD_RICJA	Q52764 rickettsia
4	285	21.3	159	1 17KD_RICCN	P05372 rickettsia
5	276.5	20.7	159	1 17KD_RICTY	P22882 rickettsia
6	268	20.1	154	1 17KD_RICAU	P50928 rickettsia
7	259	19.4	154	1 17KD_RICPA	P50930 rickettsia
8	259	19.4	154	1 17KD_RICRH	P50931 rickettsia
9	256	19.2	154	1 17KD_RICMO	P50929 rickettsia
10	255	19.1	154	1 17KD_RICAM	P50927 rickettsia
11	245	18.4	772	1 CIPB_CLOTM	Q01866 clostridium
12	245	18.4	1853	1 CIPA_CLOTM	Q06851 clostridium
13	188.5	14.1	879	1 GUNI_CLOTM	Q02934 clostridium
14	188	14.1	986	1 GUNZ_CLOSR	P23659 clostridium
15	162.5	12.2	914	1 GUX2_CLOSR	P50900 clostridium
16	156	11.7	80	1 17KD_RICCA	P29697 rickettsia
17	143.5	10.7	1039	1 GUNB_CALSA	P10474 c endoglucanase
18	143.5	10.7	1331	1 MANB_CALSA	P22533 caldocellum
19	138.5	10.4	1742	1 GUNA_CALSA	P22534 caldocellum
20	121	9.1	499	1 GUN2_BACSU	P10475 bacillus su
21	115	8.6	499	1 GUN3_BACSU	P23549 bacillus su
22	112.5	8.4	155	1 PCP_YREN	P31484 versinia en
23	111	8.3	145	1 YCEA_PAECLA	P29718 paenibacill
24	109.5	8.2	499	1 GUNI_BACSU	P07983 bacillus su
25	109	8.2	155	1 SLVB_SALTY	Q53549 salmonella
26	104.5	7.8	155	1 SLVB_ECOLI	P55741 escherichia
27	103	7.7	504	1 GUNW_ERWCA	Q59395 erwinia car
28	101.5	7.6	155	1 PCP_HAEIN	P10325 haemophilus
29	101.5	7.6	505	1 GUNV_ERWCA	Q47096 erwinia car
30	100.5	7.5	700	1 GUNA_PAECLA	P29719 paenibacill
31	100	7.5	179	1 YCFJ_ECOLI	P37796 escherichia
32	100	7.5	444	1 GUNN_ERWCA	Q59394 erwinia car
33	98	7.3	922	1 PMP1_CHLPN	Q92955 chlamydia p

ALIGNMENTS

RESULT 1

ID	CBPA_CLOCL	STANDARD;	PRT;	1848 AA.
AC	P38058;			
DT	01-OCT-1994 (Rel. 30, Created)			
DT	01-OCT-1994 (Rel. 30, Last sequence update)			
DT	15-DEC-1998 (Rel. 37, Last annotation update)			
DE	Cellulose binding protein A precursor.			
GN	CBPA.			
OS	Clostridium cellulovorans.			
OC	Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;			
OC	Clostridium.			
OX	NCBI_TaxID=1493;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92228810; PubMed=1565642;			
RA	Shoseyov O., Takagi M., Goldstein M.A., Doi R.H.;			
RT	"Primary sequence analysis of Clostridium cellulovorans cellulose			
RT	binding protein A.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 89:3483-3487 (1992).			
CC	-!- FUNCTION: BINDS TO CELLULOSE FIBERS AND COORDINATES CELLULOSE			
CC	ENZYMES.			
CC	-!- PTM: THE N-TERMINUS IS BLOCKED.			
CC	-!- PTM: CONTAINS ABOUT 10% OF CARBOHYDRATES.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	EMBL; M73817; AAA23218.1; -.			
DR	PIR; A44140; A44140.			
DR	HSP; Q06851; INBC.			
DR	InterPro; IPR001956; CBD_3.			
DR	InterPro; IPR002102; Cohesin.			
DR	InterPro; IPR005102; DUF291.			
DR	Pfam; PF00942; CBM_3; 1.			
DR	Pfam; PF00963; Cohesin; 9.			
DR	Pfam; PF03442; DUF291; 4.			
DR	ProDom; PD001947; CBD_3; 1.			
KW	Cellulose degradation; Cell			
FT	SIGNAL 1 28			
FT	CHAIN 29 1848			
FT	DOMAIN 29 189			
FT	SEQUENCE 1848 AA; 189152 MW; 85FA6CE6F771AF1A CRC64;			
SQ				
Query Match	36.5%;	Score 487;	DB 1;	Length 1848;
Best Local Similarity	46.8%;	Pred. No. 1.9e-32;		
Matches 116;	Conservative 12;	Mismatches 30;	Indels 90;	Gaps 7;
QY	1 MSVEFYNSKSAQTNSITPIIKITNTSDNLNDLVKRYYYTSDGTQGTFCWDHAGAL 60			
DB	33 MSVEFYNSKSAQTNSITPIIKITNTSDNLNDLVKRYYYTSDGTQGTFCWDHAGAL 92			

QY 61 LGNSYDNTSKVTANFVKETASPTSTYDTYLDPSHMERGCLQGSSLLIISVFLVGCQNFPS 120
 |||||
 Db 93 LGNSYDNTSKVTANFVKETASPTSTYDTYLDPSHMERGCLQGSSLLIISVFLVGCQNFPS 124
 |||||
 QY 121 ROEVAATGAVGVAGQVLFKSGSRVSMAGVILGGLIGSKIGOSMDQDQKIKLNQSL 180
 |||||
 Db 125 -----FGPASGRATL-----KKGQFIIQ----- 143
 QY 181 EKVAGQVTR--WRNPDTGNSYSVEPVRTYQVYRNKERRQOYCREFOQKAMIAQOKEIY 238
 |||||
 Db 144 -----GRITKDSNSYNTQTDNSYFDASSSTPVVNP-----KVTGYIGGAK--VL 185
 |||||
 QY 239 GTACPOPD 246
 |||||
 Db 186 GTA-PGPD 192
 |||||

RESULT 2

17KD RICPR STANDARD; PRT; 159 AA.
 ID 17KD RICPR AC P16624;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 17 kDa surface antigen precursor.
 GN OMP OR RP833.
 OS Rickettsia prowazekii.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsiae; Rickettsia.
 OX NCBI_TaxID=782;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Madrid E;
 RX MEDLINE=89359171; PubMed=2768201;
 RA Anderson B.E., Tzianabos T.;
 RT "Comparative sequence analysis of a genus-common rickettsial antigen
 gene.";
 RL J. Bacteriol. 171:5199-5201(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Madrid E;
 RX MEDLINE=99039499; PubMed=9823893;
 RA Andersson S.G.B., Zomrodipour A., Andersson J.O.,
 RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
 RT "The genome sequence of Rickettsia prowazekii and the origin of
 mitochondria.";
 RL Nature 396:133-140(1998).
 CC -!- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
 anchor (Probable).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 or send an email to license@isb-sib.ch).

Query Match 21.6%; Score 288; DB 1; Length 159;
 Best Local Similarity 37.9%; Pred. No. 3.9e-17;
 Matches 61; Conservative 32; Mismatches 52; Indels 16; Gaps 6;
 QY 61 LGNSYDNTSKVTANFVKETASPTSTYDTYLDPSHMERGCLQGSSLLIISVFLVGCQNFPS 120
 |||||
 Db 93 LGNSYDNTSKVTANFVKETASPTSTYDTYLDPSHMERGCLQGSSLLIISVFLVGCQNFPS 124
 |||||
 QY 121 ROEVAATGAVGVAGQVLFKSGSRVSMAGVILGGLIGSKIGOSMDQDQKIKLNQSL 180
 |||||
 Db 125 -----FGPASGRATL-----KKGQFIIQ----- 143
 QY 181 EKVAGQVTR--WRNPDTGNSYSVEPVRTYQVYRNKERRQOYCREFOQKAMIAQOKEIY 238
 |||||
 Db 144 -----GRITKDSNSYNTQTDNSYFDASSSTPVVNP-----KVTGYIGGAK--VL 185
 |||||
 QY 239 GTACPOPD 246
 |||||
 Db 186 GTA-PGPD 192
 |||||

Query Match 21.6%; Score 288; DB 1; Length 159;
 Best Local Similarity 37.9%; Pred. No. 3.9e-17;
 Matches 61; Conservative 32; Mismatches 52; Indels 16; Gaps 6;

QY 103 SLLIISV---FLVGC--AQNFSRQEVGAATGAVGVAGQVLFKSGSRVSMAGI-GAVL 156
 |||||
 Db 5 SKIMIILAAASMLQACNGQGMNKQGTGTLTGAGGAGALLGSGQFGQKGQL-VGVGVGALL 63
 |||||
 QY 157 GGLIGSKIGOSMDQDQK-----IKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQVYN 212
 |||||
 Db 64 GAVLGGQIGASMDQDRRLLELTSORALESPAGSGNSIENWRNPNGNHHGYVTPNKTY----- 119
 |||||
 QY 213 KOERROQYCREFOQKAMIAQOKEIYGTACPDGRWQVIS 253
 |||||
 Db 120 -RNSAGQYCREYTTQVILGGKQKQYGNACRQPDGQWQVYN 159
 |||||

RESULT 3

17KD RICJA STANDARD; PRT; 159 AA.
 ID 17KD RICJA AC Q52764;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE 17 kDa surface antigen precursor.
 GN OMP.
 OS Rickettsia japonica.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsiae; Rickettsia.
 OX NCBI_TaxID=35790;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=YH;
 RX MEDLINE=95229950; PubMed=7714214;
 RA Furuya Y., Katayama T., Yoshida Y., Kaiho I.;
 RT "Specific amplification of Rickettsia japonica DNA from clinical
 specimens by PCR.";
 RL J. Clin. Microbiol. 33:487-489(1995).
 CC -!- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
 anchor (Probable).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See http://www.isb-sib.ch/announce/
 or send an email to license@isb-sib.ch).

EMBL; D16515; BAA03965.1; -;
 DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.
 KW Outer membrane; Lipoprotein; Antigen; Signal.
 FT SIGNAL 1 19 BY SIMILARITY.
 FT CHAIN 20 159 17 kDa SURFACE ANTIGEN.
 FT LIPID 20 20 N-ACYL DIGLYCERIDE (PROBABLE).
 SQ SEQUENCE 159 AA; 16554 MW; 16554 MW; CDDCE7CEBDCD6B41 CRC64;

Query Match 21.4%; Score 286; DB 1; Length 159;
 Best Local Similarity 37.9%; Pred. No. 5.8e-17;
 Matches 61; Conservative 32; Mismatches 52; Indels 16; Gaps 6;

QY 103 SLLIISV---FLVGC--AQNFSRQEVGAATGAVGVAGQVLFKSGSRVSMAGI-GAVL 156
 |||||
 Db 5 SKIMIILAAASMLQACNGQGMNKQGTGTLTGAGGAGALLGSGQFGQKGQL-VGVGVGALL 63
 |||||
 QY 157 GGLIGSKIGOSMDQDQK-----IKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQVYN 212
 |||||
 Db 64 GAVLGGQIGASMDQDRRLLELTSORALESPAGSGNSIENWRNPNGNHHGYVTPNKTY----- 119
 |||||
 QY 213 KOERROQYCREFOQKAMIAQOKEIYGTACPDGRWQVIS 253
 |||||
 Db 120 -RNSAGQYCREYTTQVILGGKQKQYGNACRQPDGQWQVYN 159
 |||||

RESULT 4

17KD RICCN

QY 235 QEYGTACPODGRQVVS 253
 Db 141 QTYGNACRPDQGWQVN 159

RESULT 6

ID 17KD RICAU STANDARD; PRT; 154 AA.
 AC P50928;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE 17 kDa surface antigen precursor (Fragment).
 GN OMP.
 OS Rickettsia australis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsiae; Rickettsia.
 OX NCBI_TaxID=787;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Baird R.W., Ross B., Dwyer B.;
 RL Submitted (OCT-1991) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid anchor (Probable).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M74042; AAA26394.1; --
 DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.
 KW Outer membrane; Lipoprotein; Antigen; Signal.
 FT SIGNAL 1 19 BY SIMILARITY.
 FT CHAIN 20 >154 17 kDa SURFACE ANTIGEN.
 FT LIPID 20 20 N-ACYL DIGLYCERIDE (PROBABLE).
 FT NON_TER 154 154
 FT SEQUENCE 154 AA; 15967 MW; E3AA83346FAC320 CRC64;
 SQ

Query Match 20.1%; Score 268; DB 1; Length 154;
 Best Local Similarity 37.8%; Pred. No. 1.7e-15;
 Matches 59; Conservative 29; Mismatches 52; Indels 16; Gaps 6;
 QY 103 SSLIIISV---FLVGCQ--NFSRQEVGAATGAVGVGAGQLFGKSGRVSMAIG-GAVL 156
 Db 5 SKIMIIAALASMLQACNSPGMKNKGTGTLGGAGALLGSGQFGKGGQL-VGVGVGALL 63
 QY 157 GGLIGSKIGSQMDQDK---IKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYN 212
 Db 64 GAVLGGQIGAGMDEQDRRLAELTSQRALETAPSGSNVEWRNPNGNGYVTPNKTYRNSN 123
 QY 213 KQERQQYCREFOQKAMIAQKQEIYGTACPOPDGR 248
 Db 124 -----GQYCREYQTIVVIGGKQKAYGNACRPDQ 154

RESULT 7

ID 17KD RICPA STANDARD; PRT; 154 AA.
 AC P50930;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE 17 kDa surface antigen precursor (Fragment).
 GN OMP.
 OS Rickettsia parkeri.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsiae; Rickettsia.
 OX NCBI_TaxID=35792;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=Maculatum;
 RA Pretzman C.I., Stothard D.R., Ralph D., Clark J.B., Fuerst P.A.;
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid anchor (Probable).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U17008; AAA2040.1; --
 DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.
 KW Outer membrane; Lipoprotein; Antigen; Signal.
 FT SIGNAL 1 19 BY SIMILARITY.
 FT CHAIN 20 >154 17 kDa SURFACE ANTIGEN.
 FT LIPID 20 20 N-ACYL DIGLYCERIDE (PROBABLE).
 FT NON_TER 154 154
 FT SEQUENCE 154 AA; 15897 MW; 5D06F45F9DBD5EEC CRC64;
 SQ
 Query Match 19.4%; Score 259; DB 1; Length 154;
 Best Local Similarity 36.5%; Pred. No. 9.8e-15;
 Matches 57; Conservative 30; Mismatches 53; Indels 16; Gaps 6;
 QY 103 SSLIIISV---FLVGC--AQNFSRQEVGAATGAVGVGAGQLFGKSGRVSMAIG-GAVL 156
 Db 5 SKIMVIALATSMQACNGPGMKNKGTGTLGGAGALLGSGQFGKGGQL-VGVGVGALL 63
 QY 157 GGLIGSKIGSQMDQDK---IKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYN 212
 Db 64 GAVLGGQIGAGMDEQDRRLAELTSQRALETAPSGSNVEWRNPNGNGYVTPNKTY---- 119
 QY 213 KQERQQYCREFOQKAMIAQKQEIYGTACPOPDGR 248
 Db 120 -RNSTGQYCREYQTIVVIGGKQKAYGNACLPDQ 154
 RESULT 8
 ID 17KD RICRH STANDARD; PRT; 154 AA.
 AC P50931;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE 17 kDa surface antigen precursor (Fragment).
 GN OMP.
 OS Rickettsia rhipicephali.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsiae; Rickettsia.
 OX NCBI_TaxID=33992;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Stothard D.R., Ralph D.A., Clark J.B., Fuerst P.A., Pretzman C.;
 RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid anchor (Probable).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U11020; AAB07706.1; --
 DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.
 KW Outer membrane; Lipoprotein; Antigen; Signal.
 FT SIGNAL 1 19 BY SIMILARITY.


```

RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=YS;
RA MEDLINE=93146373; PubMed=1490597;
RA Poole D.M., Morag E., Lamed R., Bayer E.A., Hazlewood G.P.,
RA Gilbert H.J.;
RT subunit S1 from Clostridium thermocellum YS.;
RL FEMS Microbiol. Lett. 78:181-186(1992).
CC -!- FUNCTION: ACTS AS A SCAFFOLDING PROTEIN IN THE CELLULOSE. IT
CC PROMOTES BINDING OF CELLULOSE TO THE CATALYTIC DOMAINS OF THE
CC CELLULOXYLIC ENZYMES PROBABLY THROUGH THE BINDING OF THE NINE
CC REPEATED DOMAINS WITH THE DUPLICATED SEGMENT OF 24 AMINO ACIDS
CC PRESENT IN CATALYTIC SUBUNITS OF THE CELLULOSE.
CC -!- SUBCELLULAR LOCATION: CELL SURFACE.
CC -!- DOMAIN: THE COHESIN DOMAINS BIND TO THE DOCKERIN DOMAIN BORN BY
CC THE CATALYTIC COMPONENTS OF THE CELLULOSE.
CC -!- SIMILARITY: Contains at least 3 cohesin domains.
CC -!- SIMILARITY: Contains 2 dockerin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X68233; CAA48312.1; -.
CC HSSP; Q06851; INBC.
CC InterPro; IPR001956; Cohesin_3.
CC InterPro; IPR002102; Cohesin.
CC InterPro; IPR002105; Dockerin_1.
CC InterPro; IPR002048; EF-hand.
CC Pfam; PF00942; CBM_3; 1.
CC Pfam; PF00963; Cohesin; 3.
CC Pfam; PF00404; Dockerin_1; 2.
CC ProDom; PD001947; CBD_3; 1.
CC PROSITE; PS00018; EF_HAND; UNKNOWN 1.
CC PROSITE; PS00448; CLOS_CELLULOSE_RPT; 2.
KW Cellulose degradation; Cell wall; Glycoprotein; Repeat.
FT NON_TER 1
FT DOMAIN <1 80 COHESIN 1.
FT LINKER (PRO/THR-RICH) .
FT DOMAIN 81 93 COHESIN 2.
FT LINKER (PRO/THR-RICH) .
FT DOMAIN 94 240 COHESIN 3.
FT LINKER (PRO/THR-RICH) .
FT DOMAIN 241 272 CELLULOSE-BINDING.
FT DOMAIN 273 439 CELLULOSE-BINDING.
FT DOMAIN 440 461 LINKER (PRO/THR-RICH) .
FT DOMAIN 462 607 COHESIN 3.
FT DOMAIN 710 733 DOCKERIN 1.
FT DOMAIN 743 766 DOCKERIN 2.
SQ SEQUENCE 772 AA; 82491 MW; BBF06D5E094FE10 CRC64;

Query Match 18.4%; Score 245; DB 1; Length 772;
Best Local Similarity 51.5%; Pred. No. 9.2e-13;
Matches 53; Conservative 16; Mismatches 30; Indels 4; Gaps 3;

QY 1 MSVEFYNSKSAQTNSITPIKNTSDSLNLDVKKVRYXTSDGTGCTFWCDHAGAL 60
Db 281 LKVEFYNSPSTTNSINPQFVNTGSSAIDLSKULTRYIYVDGQDQTFWCDHA-AI 339
QY 61 LGN--SYVDNTSKVTANFKETASPTSTYDTYLDPSHMRGCLQ 101
Db 340 IGSNGSYNGITSNVKGTFVKMSSS-TNNADTVLEISFTGTGLE 381

RESULT 12
CIPA_CLOTM STANDARD; PRT; 1853 AA.
AC Q06851.
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)

```

```

DE Cellulosomal scaffolding protein A precursor (Cellulosomal
DE glycoprotein S1/SL) (Cellulose integrating protein A) (Cohesin).
DE CIPA.
OS Clostridium thermocellum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1515;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 29-40 AND 42-43.
RC STRAIN=ATCC 27405 / DSM 1237;
RX MEDLINE=93302508; PubMed=8316083;
RA Gerngross U.T., Romaniec M.P.M., Kobayashi T., Huekisson N.S.,
RA Demain A.L.;
RT "Sequencing of a Clostridium thermocellum gene (cipA) encoding the
RT cellulosomal SL-protein reveals an unusual degree of internal
RT homology.";
RL Mol. Microbiol. 8:325-334(1993).
RN [2]
RP SEQUENCE OF 1820-1853 FROM N.A.
RX MEDLINE=93209931; PubMed=8458832;
RA Fujino T., Beguin P., Aubert J.-P.;
RT "Organization of a Clostridium thermocellum gene cluster encoding the
RT cellulosomal scaffolding protein Cipa and a protein possibly involved
RT in attachment of the cellulosome to the cell surface.";
RL J. Bacteriol. 175:1891-1899(1993).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.15 ANGSTROMS) OF 184-321.
RX MEDLINE=97238934; PubMed=9083107;
RA Shimon L.J.W., Bayer E.A., Morag E., Lamed R., Yaron S., Shoham Y.,
RA Prolow F.;
RT "A cohesin domain from Clostridium thermocellum: the crystal
RT structure provides new insights into cellulosome assembly.";
RL Structure 5:381-390(1997).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 368-522.
RX MEDLINE=97076134; PubMed=8918451;
RA Torno J., Lamed R., Chirino A.J., Morag E., Bayer E.A., Shoham Y.,
RA Steitz T.A.;
RT "Crystal structure of a bacterial family-III cellulose-binding
RT domain: a general mechanism for attachment to cellulose.";
RL EMBO J. 15:5739-5751(1996).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 1213-1361.
RX MEDLINE=98022914; PubMed=9402065;
RA Tavares G.A., Beguin P., Alzari P.M.;
RT "The crystal structure of a type I cohesin domain at 1.7-A
RT resolution.";
RL J. Mol. Biol. 273:701-713(1997).
CC -!- FUNCTION: ACTS AS A SCAFFOLDING PROTEIN IN THE CELLULOSE. IT
CC PROMOTES BINDING OF CELLULOSE TO THE CATALYTIC DOMAINS OF THE
CC CELLULOXYLIC ENZYMES.
CC -!- SUBCELLULAR LOCATION: CELL SURFACE.
CC -!- DOMAIN: THE COHESIN DOMAINS BIND TO THE DOCKERIN DOMAIN BORN BY
CC THE CATALYTIC COMPONENTS OF THE CELLULOSE.
CC -!- SIMILARITY: Contains 9 cohesin domains.
CC -!- SIMILARITY: Contains 2 dockerin domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L08665; -. NOT ANNOTATED CDS.
CC EMBL; X67506; CAA47840.1; -.
CC PIR; S36859; S36859.
CC PDB; 1ANU; 23-JUL-97.
CC PDB; 1AOH; 08-JUL-98.
CC PDB; INBC; 26-SEP-97.
CC InterPro; IPR001956; CBD_3.
CC InterPro; IPR002102; Cohesin.

```


DR InterPro; IPR002105; Dockerin_1.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00942; CEM 3; 1.
DR Pfam; PF00963; Cohesin; 9.
DR Pfam; PF00404; Dockerin_1; 2.
DR ProDom; PD001947; CBD3; 1.
DR PROSITE; PS00018; EF_HAND; UNKNOWN 1.
DR PROSITE; PS00448; CLOS_CELLULOSE_RPT; 2.
KW Cellulose degradation; Cell wall; Glycoprotein; Repeat; Signal;
3D-structure.
FT SIGNAL 1 28
FT CHAIN 29 1853
FT DOMAIN 29 182
FT DOMAIN 183 322
FT DOMAIN 323 363
FT DOMAIN 364 522
FT DOMAIN 523 559
FT DOMAIN 560 704
FT DOMAIN 724 866
FT DOMAIN 889 1031
FT DOMAIN 1054 1196
FT DOMAIN 1219 1361
FT DOMAIN 1384 1526
FT DOMAIN 1548 1690
FT DOMAIN 1791 1814
FT DOMAIN 1824 1847
FT CONFLICT 1615 1615
FT STRAND 185 188
FT STRAND 190 191
FT TURN 195 196
FT STRAND 198 206
FT TURN 210 211
FT STRAND 213 221
FT TURN 224 226
FT STRAND 227 234
FT TURN 236 237
FT TURN 243 246
FT STRAND 247 252
FT TURN 253 256
FT STRAND 257 263
FT TURN 265 266
FT TURN 270 271
FT STRAND 273 273
FT STRAND 277 286
FT STRAND 292 304
FT TURN 305 306
FT STRAND 309 309
FT STRAND 313 315
FT STRAND 317 319
FT STRAND 369 375
FT STRAND 381 382
FT STRAND 385 385
FT STRAND 387 393
FT STRAND 399 400
FT HELIX 401 403
FT STRAND 404 410
FT STRAND 418 428
FT TURN 430 431
FT STRAND 434 436
FT HELIX 438 440
FT STRAND 441 452
FT TURN 453 454
FT STRAND 455 463
FT STRAND 467 468
FT TURN 470 471
FT STRAND 473 482
FT TURN 483 484
FT STRAND 488 489
FT TURN 491 492
FT TURN 494 495
FT STRAND 498 498
FT STRAND 503 504
FT STRAND 509 512

FT TURN 513 514
FT STRAND 515 518
FT STRAND 1220 1224
FT STRAND 1226 1229
FT TURN 1231 1232
FT STRAND 1234 1242
FT TURN 1246 1247
FT STRAND 1249 1249
FT STRAND 1251 1257
FT TURN 1260 1262
FT STRAND 1263 1270
FT TURN 1272 1273
FT HELIX 1279 1282
FT STRAND 1283 1288
FT TURN 1289 1292
FT STRAND 1293 1299
FT TURN 1301 1302
FT TURN 1306 1307
FT STRAND 1309 1309
FT STRAND 1313 1322
FT TURN 1324 1325
FT STRAND 1329 1342
FT TURN 1344 1345
FT STRAND 1348 1348
FT STRAND 1351 1354
FT STRAND 1356 1360
SQ SEQUENCE 1853 AA; 3ABDDC03ABFC5372 CRC64;
Query Match 18.4%; Score 245; DB 1; Length 1853;
Best Local Similarity 51.5%; Pred. No. 2.5e-12;
Matches 53; Conservative 16; Mismatches 30; Indels 4; Gaps 3;
QY 1 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDNVKRYVYTSDTGTGQTFWCDHAGAL 60
Db 369 LKVEFYNSPDDTNSINPQPKVTNTGSSADLSKLTLYRYTVGQKDTFWCDHA-AI 427
QY 61 LGN--SYVDNTSKVTANFVKETASPTSTYDLPDPSHMRGCIQ 101
Db 428 IGSNGSYNGITSNVKGTFFVKMSSS-TNNADTYLEISFTGGTLE 469
RESULT 13
GUNI_CLOTM STANDARD; PRT; 879 AA.
AC Q02934;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Endoglucanase I precursor (EC 3.2.1.4) (EGI) (Endo-1,4-beta-glucanase)
DE (Cellulase I).
GN CELL.
OS Clostridium thermocellum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1515;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 56-69.
RC STRAIN=NCIB 10682;
RX MEDLINE=93171873; PubMed=8436949;
RA Hazlewood G.P., Davidson K., Laurie J.I., Huskisson N.S.,
RA Gilbert H.J.;
RT "Gene sequence and properties of Cell, a family E endoglucanase from
Clostridium thermocellum.";
RL J. Gen. Microbiol. 139:307-316(1993).
CC -|- FUNCTION: THIS ENZYME CATALYZES THE ENDOHYDROLYSIS OF 1,4-BETA-
GLUCOSIDIC LINKAGES IN CELLULOSE, LICHENIN AND CEREAL BETA-D-
GLUCANS. PRINCIPALLY ACTIVE AGAINST BARLEY BETA-GLUCAN.
CC -|- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-glucosidic
linkages in cellulose, lichenin and cereal beta-D-glucans.
CC -|- PATHWAY: Cellulose degradation.
CC -|- SIMILARITY: BELONGS TO CELLULASE FAMILY E (FAMILY 9 OF GLYCOSYL
HYDROLASES).
CC -----

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 20:09:46 ; Search time 37.582 Seconds
(without alignments)
1757.794 Million cell updates/sec

Title: US-09-677-374-6
Perfect score: 1335
Sequence: 1 MSVRFYNSKSAQTNSITPI.....IYGTACQPQDGRWQVISTEK 256

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_23.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	815	61.0	2 Q9F9K8	Q9F9K8 pisciricket
2	303.5	22.7	148 2 O54381	O54381 rickettsia
3	283	21.2	159 2 Q9F9F2	Q9F9F2 rickettsia
4	258.5	19.4	137 2 O52252	O52252 rickettsia
5	252.5	18.9	144 2 Q9K2N6	Q9K2N6 male-killin
6	251.5	18.8	137 2 O31065	O31065 rickettsia
7	251.5	18.8	144 2 Q9K4W8	Q9K4W8 male-killin
8	249	18.7	1546 2 Q45996	Q45996 clostridium
9	244	18.3	154 2 Q53154	Q53154 rickettsia
10	241.5	18.1	131 2 Q8KLT4	Q8KLT4 rickettsia
11	241.5	18.1	131 2 Q8KLT5	Q8KLT5 rickettsia
12	239	17.9	151 2 Q9F9Q9	Q9F9Q9 rickettsia
13	236.5	17.7	131 2 Q9F0Q1	Q9F0Q1 rickettsia
14	236.5	17.7	131 2 Q52637	Q52637 rickettsia
15	235.5	17.6	131 2 Q8KLU2	Q8KLU2 rickettsia
16	235.5	17.6	131 2 Q9L522	Q9L522 rickettsia

17	233.5	17.5	1162	2	O82830	O82830 clostridium
18	223.5	16.7	1483	16	Q977Y4	Q977Y4 clostridium
19	216.5	16.2	105	2	O31208	O31208 rickettsia
20	195.5	14.6	1915	2	Q9RPL0	Q9RPL0 acetivibrio
21	193	14.5	887	2	Q9L3J8	Q9L3J8 clostridium
22	172.5	12.9	307	2	Q46392	Q46392 clostridium
23	172	12.9	1091	2	Q8KKF7	Q8KKF7 paenibacill
24	152.5	11.4	997	2	Q924I1	Q924I1 bacillus sp
25	151.5	11.3	1779	2	O52374	O52374 caldicellul
26	150.5	11.3	1770	2	Q9X3P5	Q9X3P5 caldicellul
27	149.5	11.2	261	2	Q9AQG7	Q9AQG7 caldicellul
28	149.5	11.2	1426	2	Q9X3P6	Q9X3P6 caldicellul
29	149	11.2	221	2	Q8VVI7	Q8VVI7 clostridium
30	148.5	11.1	1751	2	Q9AQG4	Q9AQG4 caldicellul
31	144.5	10.8	996	2	Q9AQH0	Q9AQH0 caldicellul
32	142.5	10.7	1000	2	O24820	O24820 thermophili
33	142	10.6	1711	2	P96311	P96311 anaerocellu
34	139	10.4	77	2	Q9AGC7	Q9AGC7 rickettsia
35	138.5	10.4	921	2	Q9L8L8	Q9L8L8 caldicellul
36	137	10.3	199	16	Q985G4	Q985G4 rhizobium 1
37	135.5	10.1	930	2	Q9RFX5	Q9RFX5 caldicellul
38	129.5	9.7	161	16	Q8GIU7	Q8GIU7 brucella su
39	128	9.6	170	2	Q9RFX6	Q9RFX6 caldicellul
40	127.5	9.6	182	16	Q9HXI3	Q9HXI3 pseudomonas
41	126.5	9.5	135	2	Q8VUE8	Q8VUE8 brucella ab
42	122	9.1	257	16	Q9A8M8	Q9A8M8 caulobacter
43	121	9.1	499	2	Q93TJ6	Q93TJ6 bacillus su
44	121	9.1	508	2	Q93LD0	Q93LD0 bacillus su
45	120.5	9.0	131	16	Q8YQZ7	Q8YQZ7 brucella me

ALIGNMENTS

RESULT 1

Q9F9K8 ID Q9F9K8 PRELIMINARY; PRT; 162 AA.
AC Q9F9K8; DT 01-MAR-2001 (TRENBLrel. 16, Created)
RC STRAIN=LF-89; DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE 17 kDa antigen.
GN OSPA.
OS Piscirickettsia salmonis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Thiotrichales;
OC Piscirickettsiaceae; Piscirickettsia.
OX NCBI_taxid=1238;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LF-89;
RA Kuzyk M.A., Burian J., Thornton J.C., Kay W.W.;
RT "Identification of a genus-common Rickettsial surface antigen in the salmonid pathogen Piscirickettsia salmonis."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF184152; AAG17000.1; -
SQ SEQUENCE 162 AA; 17661 MW; DDE9956FD94A527E CRC64;

Query Match 61.0%; Score 815; DB 2; Length 162;
Best Local Similarity 98.8%; Pred. No. 3.9e-60;
Matches 158; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 97 RGCIGSSLLIISVFLVGCANFSEQVEGATGAVVGVAGOLFQKSGRVSMAIGGAVL 156
Db 3 RGCIGSSLLIISVFLVGCANFSEQVEGATGAVVGVAGOLFQKSGRVSMAIGGAVL 62
QY 157 GGLIGSKTGQSDQDDKIKLNQSLKVKAGQVTRWRNPDTGNSYSEVPVRYQYKQER 216
Db 63 GGLIGSKTGQSDQDDKIKLNQSLKVKAGQVTRWRNPDTGNSYSEVPVRYQYKQER 122
QY 217 RQYCREPQQAMTIAQKQEIYGTACPPDGRWQVISTEK 256
Db 123 RQYCREPQQAMTIAQKQEIYGTACPPDGRWQVISTEK 162

RESULT 2
 OS4381 PRELIMINARY; PRT; 148 AA.
 ID O54381
 AC O54381
 DT 01-JUN-1998 (TReMBLrel. 06, Created)
 DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
 DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
 DE 17 kDa common-antigen (Fragment).
 OS Rickettsia sp.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsiae; Rickettsia.
 OX NCBI_TaxID=789;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98087556; PubMed=9425244;
 RA Davis M.J., Ying Z., Brunner B.R., Pantoja A., Ferwerda F.H.;
 RT "Rickettsial relative associated with papaya bunchy top disease."
 RL Curr. Microbiol. 36:80-84(1998).
 DR EMBL; U76907; AAC02809.1; -;
 FT NON_TER 1
 FT 148
 SQ SEQUENCE 148 AA; 15050 MW; A7AFEEFDE0AEE4C CRC64;
 Query Match 22.7%; Score 303.5; DB 2; Length 148;
 Best Local Similarity 40.1%; Pred. No. 1.3e-17;
 Matches 55; Conservative 30; Mismatches 43; Indels 9; Gaps 2;
 QY 120 SRQEVGAATGAVVGVAGQLFGKSGRVSMAIGAVLGLGSKIGSQMDQDK----IK 175
 Db 17 NKQSGTLLGTLGGVSGFGGTLAAVAGALLGAILNQIGAGMDEQDRKLAELT 76
 QY 176 LNQSLKVKVAGQVTRWRNPDGNSYSVEPVRYQYKQRRQYCRFQKAMTAGQKQ 235
 Db 77 SQRALEAPSGSVWRNPDGNGYGTPTSKAY-----KNNTGQYCREYTTQTVVVGKQ 131
 QY 236 EYGTACPODGRWQVI 252
 Db 132 KAYGTACRQPDQWQV 148
 RESULT 3
 Q9F9F2 PRELIMINARY; PRT; 159 AA.
 ID Q9F9F2
 AC Q9F9F2
 DT 01-MAR-2001 (TReMBLrel. 16, Created)
 DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
 DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
 DE 17 kDa genus-common antigen.
 OS Rickettsia felis (Rickettsia azadi).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsiae; Rickettsia.
 OX NCBI_TaxID=42862;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21217364; PubMed=11321078;
 RA Bouyer D.H., Stenos J., Crocquet-Valdes P., Moron C.G., Popov V.I.,
 RA Zavala-Velazquez J.E., Foil L.D., Stochard D.R., Azad A.F.,
 RA Walker D.H.;
 RT "Rickettsia felis: molecular characterization of a new member of the
 RT spotted fever group."
 RL Int. J. Syst. Evol. Microbiol. 51:339-347(2001).
 DR EMBL; AF195118; AG28452.1; -;
 SQ SEQUENCE 159 AA; 16497 MW; 34C5B020AF470A1F CRC64;
 Query Match 21.2%; Score 283; DB 2; Length 159;
 Best Local Similarity 37.9%; Pred. No. 7e-16;
 Matches 61; Conservative 30; Mismatches 54; Indels 16; Gaps 6;
 QY 103 SLLIISV---FLVGC--AQNFSRQEVGAATGAVVGVAGQLFGKSGRVSMAIG-GAVL 156
 Db 5 SKIMIIAALASMLQACNPGGMNKQGTLLGGAGGALLGSGFGKQQL-VGVGVGALL 63

QY 157 GGLIGSKIGSQMDQDK----IKLQSLKVKVAGQVTRWRNPDGNSYSVEPVRYQYK 212
 Db 64 GAVLGGQIGAGMDEQDRKLAELTSQRALEAPSGTSVEWRNPDGNGHGYVTNKTY---- 119
 QY 213 KOERRQYCRFQKAMTAGQKQEIYGTACPODGRWQVIS 253
 Db 120 -RNSTGQYCREYTTQTVVVGKQKAYGNACRQPDGLWQVWN 159
 RESULT 4
 OS2252 PRELIMINARY; PRT; 137 AA.
 ID OS2252
 AC OS2252
 DT 01-JUN-1998 (TReMBLrel. 06, Created)
 DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
 DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
 DE 17 kDa antigen (Fragment).
 OS Rickettsia cooley.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsiae; Rickettsia.
 OX NCBI_TaxID=69410;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Billings A.N., Teltow G.J., Walker D.H.;
 RT "Molecular characterization of a novel spotted fever group rickettsial
 RT species from ixodes scapularis in Texas."
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF031534; AAB95267.1; -;
 FT NON_TER 1
 FT 137
 SQ SEQUENCE 137 AA; 14215 MW; A27597A9AFD85FC3 CRC64;
 Query Match 19.4%; Score 258.5; DB 2; Length 137;
 Best Local Similarity 39.3%; Pred. No. 6.3e-14;
 Matches 53; Conservative 25; Mismatches 46; Indels 11; Gaps 4;
 QY 117 QNFSRQEVGAATGAVVGVAGQLFGKSGRVSMAIG-GAVLGLGSKIGSQMDQDK-- 173
 Db 7 RGMNKQGTGTLGGAGGALLGSGFGKQQL-VGVGVGALLGAVLGGQIGAGMDEQDRRL 65
 QY 174 --IKLQSLKVKVAGQVTRWRNPDGNSYSVEPVRYQYKQRRQYCRFQKAMIA 231
 Db 66 AELTSQRALEAPSGSSTWRNPDGNGYGTPTNKTY-----RNSTGQYCREYTTQTVVIG 120
 QY 232 GQKQIYGTACPOPD 246
 Db 121 GQKQKAYGNACRQPD 135
 RESULT 5
 Q9K2N6 PRELIMINARY; PRT; 144 AA.
 ID Q9K2N6
 AC Q9K2N6
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
 DE 17kDa antigen (17 kDa antigen) (Fragment).
 OS male-killing Rickettsia from Adalia bipunctata.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsiae; Rickettsia.
 OX NCBI_TaxID=38028;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Schlenker H.J.G.V.D., Habig M., Sloggett J.J., Webberley M.K.,
 RA Bertrand D., Hurst G.D.D., Majerus M.E.N.;
 RT "On the evolution of male-killing: Monophyletic origin and horizontal
 RT transfer of male-killing Rickettsia (a-Proteobacteria) from two con-
 RT generic ladybirds, Adalia bipunctata L. and A. decempunctata L.
 RT (Coleoptera: Coccinellidae)."
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ269518; CAB96383.1; -;
 DR EMBL; AJ269517; CAB96382.1; -;
 FT NON_TER 1
 FT 144

17 kDa antigen (Fragment).

OS male-killing Rickettsia from *Adalia decempunctata*.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
NCBI_TaxID=120393;
[1]
RN
RP SEQUENCE FROM N.A.
MEDLINE=20575219; PubMed=11133455;
RA Schlenker H.J.G.V.D., Habig M., Sloggett J.J., Webberley M.K.,
RA Bertrand D., Hurst G.D.D., Majerus M.E.N.;
RT "Incidence of male-killing Rickettsia spp. (alpha-Proteobacteria) in
RT the Ten-Spot Ladybird Beetle *Adalia decempunctata* L. (Coleoptera:
RT Coccinellidae).";
RT Appl. Environ. Microbiol. 67:270-277 (2001).
DR EMBL; AJ269516; CAB96381.1; -.
FT NON TER 1
FT NON TER 144 144
FT SEQUENCE 144 AA; 14801 MW; C825472F16A56AE7 CRC64;
SQ

Query Match 18.8%; Score 251.5; DB 2; Length 144;
Best Local Similarity 39.4%; Pred. No. 2.6e-13;
Matches 52; Conservative 24; Mismatches 45; Indels 11; Gaps 4

QY 120 SRQEVGATGAVVGGVAGQLFGKSGRVSNAG-
DB 17 NKQGTGTLGGAGALLGSGFGKGGQL-VGVG
QY 175. KLNGLSEKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKERRQQYCRFQOKAMTAGOK
DB 76 TQRALEAPSGSNVWRPNPDNGHGVYTNKY-----RNSTGYCREYQTQVWIGGKQ
QY 235 QEYGTACQPPD 246
DB 131 QKSYGNACRQPD 142

RESULT 8
Q45996 PRELIMINARY; PRT; 1546 AA.
Q45996;
AC
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-WAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-WAR-2003 (TREMBLrel. 23, Last annotation update)
OC Scaffolding protein precursor.
OS Clostridium cellulolyticum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
NCBI_TaxID=1521;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35319;
RX MEDLINE=96218696; PubMed=8636029;
RA Pages S., Belaich A., Tardif C., Reverbel-Leroy C., Gaudin C.,
RA Belaich J.P.;
RT "Interaction between the endoglucanase Cella and the scaffolding
RT protein CipC of the Clostridium cellulolyticum cellulosome.";
RL J. Bacteriol. 178:2279-2286(1996).
[2]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35319;
RX MEDLINE=99173902; PubMed=10074072;
RA Pages S., Belaich A., Fierobe H.P., Tardif C., Gaudin C.,
RA Belaich J.P.;
RT "Sequence analysis of scaffolding protein CipC and ORFXp, a new
RT cohesin-containing protein in Clostridium cellulolyticum: comparison
RT of various cohesin domains and subcellular localization of ORFXp.";
RL J. Bacteriol. 181:1801-1810(1999).
[3]
RN
RP
RC STRAIN=ATCC 35319;
RX Reverbel-Leroy C., Tardif C., Belaich A., Bernadac A., Gaudin C.,
RA Belaich J.

17 kDa antigen (Fragment).

OS male-killing Rickettsia from *Adalia decempunctata*.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
NCBI_TaxID=120393;
[1]
RN
RP SEQUENCE FROM N.A.
MEDLINE=20575219; PubMed=11133455;
RA Schlenker H.J.G.V.D., Habig M., Sloggett J.J., Webberley M.K.,
RA Bertrand D., Hurst G.D.D., Majerus M.E.N.;
RT "Incidence of male-killing Rickettsia spp. (alpha-Proteobacteria) in
RT the Ten-Spot Ladybird Beetle *Adalia decempunctata* L. (Coleoptera:
RT Coccinellidae).";
RT Appl. Environ. Microbiol. 67:270-277 (2001).
DR EMBL; AJ269516; CAB96381.1; -.
FT NON TER 1
FT NON TER 144 144
FT SEQUENCE 144 AA; 14801 MW; C825472F16A56AE7 CRC64;
SQ

Query Match 18.8%; Score 251.5; DB 2; Length 144;
Best Local Similarity 39.4%; Pred. No. 2.6e-13;
Matches 52; Conservative 24; Mismatches 45; Indels 11; Gaps 4

QY 120 SRQEVGATGAVVGGVAGQLFGKSGRVSNAG-
DB 17 NKQGTGTLGGAGALLGSGFGKGGQL-VGVG
QY 175. KLNGLSEKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKERRQQYCRFQOKAMTAGOK
DB 76 TQRALEAPSGSNVWRPNPDNGHGVYTNKY-----RNSTGYCREYQTQVWIGGKQ
QY 235 QEYGTACQPPD 246
DB 131 QKSYGNACRQPD 142

RESULT 8
Q45996 PRELIMINARY; PRT; 1546 AA.
Q45996;
AC
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-WAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-WAR-2003 (TRENBLrel. 23, Last annotation update)
DB SCafolding protein precursor.
OS Clostridium cellulolyticum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
NCBI_TaxID=1521;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35319;
RX MEDLINE=96218696; PubMed=8636029;
RA Pages S., Belaich A., Tardif C., Reverbel-Leroy C., Gaudin C.,
RA Belaich J.P.;
RT "Interaction between the endoglucanase Cella and the scaffolding
RT protein CipC of the Clostridium cellulolyticum cellulosome.";
RL J. Bacteriol. 178:2279-2286(1996).
[2]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35319;
RX MEDLINE=99173902; PubMed=10074072;
RA Pages S., Belaich A., Fierobe H.P., Tardif C., Gaudin C.,
RA Belaich J.P.;
RT "Sequence analysis of scaffolding protein CipC and ORFXp, a new
RT cohesin-containing protein in Clostridium cellulolyticum: comparison
RT of various cohesin domains and subcellular localization of ORFXp.";
RL J. Bacteriol. 181:1801-1810(1999).
[3]
RN
RP
RC STRAIN=ATCC 35319;
RX Reverbel-Leroy C., Tardif C., Belaich A., Bernadac A., Gaudin C.,
RA Belaich J.

"Molecular study and overexpression of the Clostridium cellulolyticum celCCF cellulase gene in Escherichia coli."
Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.

RT RP SEQUENCE FROM N.A.
RL RC STRAIN=ATCC 35119;
RL RA Pages S., Belaich A., Reverbel C., Tardif C., Fierobe H.P., Gaudin C.,
RN BELAICH J.P.;
RP SUBMITTED (OCT-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; U40345; AAC28899.2; --
DR HSSP; Q06851; INEC.
DR InterPro; IPR001956; CBD_3.
DR InterPro; IPR002102; Cohesin.
DR InterPro; IPR005102; DUF291.
DR InterPro; IPR006162; Pfantne_attach.
DR Pfam; PF00942; CBM_3; 1.
DR Pfam; PF00963; Cohesin; 8.
DR Pfam; PF03442; DUF291; 2.
DR PRODOM; PD001947; CBD_3; 1.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 7.
KW SIGNAL.
FT CHAIN 1 27 POTENTIAL.
FT CHAIN 28 1546 SCAFFOLDING PROTEIN.
SQ SEQUENCE 1546 AA; 158748 MW; F8651504EC27809F CRC64;

Query Match 18.7%; Score 249; DB 2; Length 1546;
Best Local Similarity 33.8%; Pred. No. 9.4e-12;
Matches 72; Conservative 32; Mismatches 81; Indels 28; Gaps 7;

QY 1 MSVEFYNKSAQNTGIPKIKNTSDNLNDVKVRYTSDGTQGTWCDHAGAL 60
DB 33 VSVQFNSSGSPASSNSIYARFKVTNTSGSPINLADLKLRYYTQDADKPLTFWCDHAGYM 92
QY 61 LGNSYVDNTSKVTANFVKETASTPTDYTYLD---PSHMRGCLQSSLLIISVFLVGCQA 117
DB 93 SGNYSYIDATSKVGTGSF-KAVSPAVTNADHYLEVALNSDAGSLPAGGSIEIOTRFARNDS 151
QY 118 NFRSQEVGAATGAVGVAGQFQKGS---GRVSNMAIGGAVLGGILGSKIGQSMDDQDK 173
DB 152 NPDQSDNSVYTA-----GSYMDQKISAFVGGTLAYG--STPDGNGPPQDP 197
QY 174 IKLQSLKVKAG-----QVTRWRNPDTGNSYS 201
DB 198 -TINPTTSISAKGSPADTKITLTPNGNTFNGIS 229

RESULT 9

Q53154 ID Q53154 PRELIMINARY; PRT; 154 AA.
AC Q53154;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE (Clone PRB FISF 1), 5' end CDS (Fragment).
OS Rickettsia sp.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiaseae; Rickettsia.
OX NCBI_TaxID=789;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93084757; PubMed=1452660;
RA Baird R.W., Lloyd M., Stenos J., Ross B.C., Stewart R.S., Dwyer B.;
RT "Characterization and comparison of Australian human spotted fever group rickettsiae."
RL J. Clin. Microbiol. 30:2896-2902 (1992).
DR EMBL; M99391; AAA73386.1; --
FT NON-TER 154
SQ SEQUENCE 154 AA; 15849 MW; F5C35855EDB439D2 CRC64;

Query Match 18.3%; Score 244; DB 2; Length 154;
Best Local Similarity 35.3%; Pred. No. 1.2e-12;
Matches 55; Conservative 31; Mismatches 54; Indels 16; Gaps 6;

QY 103 SSLIIISV---FLVGC--AQNFSROEVGAATGAVGVAGQFQKGSGRVSNMAIG-GAVL 156
DB 5 SKIMIIATATSMLOACNGPGGNKOGTGLTGAGAGALLGSGFGKGGQL-VGVGVGALL 63
QY 157 GGLIGSKIGQSMDDQDK-----IKLNSLEKVKAGQVTRWRNPDTGNSYSVEPRTYQRYN 212
DB 64 GAVLGGQIGAGMDEQDRRLAELTSQRALETAPSGSNVWRNPDPNGSYGYVTNKNYRST 123
QY 213 QKEREQQYCREPQOKAMTAGQKQEIYGTACPDGR 248
DB 124 QGD-----CRVYTQTVVIGGKQKAYGNACRQPDQ 154

RESULT 10

Q8KLT4 ID Q8KLT4 PRELIMINARY; PRT; 131 AA.
AC Q8KLT4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Outer membrane protein (Fragment).
GN 17 KDA.
OS Rickettsia sp. IrItA3.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiaseae; Rickettsia.
OX NCBI_TaxID=184232;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IrItA3;
RX MEDLINE=22182650; PubMed=12194779;
RA Beninati T., Lo N.L., Noda H., Esposito F., Rizzoli A., Favio G.,
RA Genchi C.;
RT "First detection of spotted fever group rickettsiae in Ixodes ricinus from Italy."
RL Emerg. Infect. Dis. 8:983-986 (2002).
DR EMBL; AJ427883; CAD20879.1; --
FT NON-TER 1
FT NON-TER 131
SQ SEQUENCE 131 AA; 13358 MW; D4152713C9FAA9CA CRC64;

Query Match 18.1%; Score 241.5; DB 2; Length 131;
Best Local Similarity 39.1%; Pred. No. 1.5e-12;
Matches 50; Conservative 24; Mismatches 43; Indels 11; Gaps 4;

QY 120 SROEVGAATGAVGVAGQFQKGSGRVSNMAIG-GAVLGGILGSKIGQSMDDQDK----I 174
DB 10 NKQGTGLTGAGAGALLGSGFGKGGQL-VGVGVGALLGAVIGGQIGAGMDEQDRRLAEL 68
QY 175 KLNOSLEKVKAGQVTRWRNPDTGNSYSVEPRTYQRYNKRQYCREPQOKAMTAGOK 234
DB 69 TSQRALEAAPSGSSTEWNPDPNGYVTPNKTY-----RNSTQYCREYQTVVIGGKQ 123
QY 235 QEYGTAC 242
DB 124 QKAYGNAC 131

RESULT 11

Q8KLT5 ID Q8KLT5 PRELIMINARY; PRT; 131 AA.
AC Q8KLT5;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Outer membrane protein (Fragment).
GN 17 KDA.
OS Rickettsia sp. IrItA2.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiaseae; Rickettsia.
OX NCBI_TaxID=184231;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IrItA2;


```

RX MEDLINE=22182650; PubMed=12194779;
RA Beninati T., Lo N.L., Noda H., Esposito F., Rizzoli A., Favia G.,
RA Genchi C.;
RT "First detection of spotted fever group rickettsiae in Ixodes ricinus
RT from Italy.";
RL Emerg. Infect. Dis. 8:983-986(2002).
DR EMBL; AJ427882; CAD20878.1; -.
FT NON TER 1
FT NON TER 131
SQ SEQUENCE 131 AA; 13358 MW; D4152713C9FAA9CA CRC64;

Query Match 18.1%; Score 241.5; DB 2; Length 131;
Best Local Similarity 39.1%; Pred. No. 1.5e-12;
Matches 50; Conservative 24; Mismatches 43; Indels 11; Gaps 4;

QY 120 SRQEVGAATGAVVGVAGQLFGKSGRVSMATG-GAVLGGLIGSKIGQSMDOODK---I 174
Db 10 NKQGTGTLGGAGGALLGSQFGKGGQL-VGVGVGALLGVGGQIGAGMDEQDRRLAEL 68

QY 175 KLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKRQRRQYQYCFQOKAMIAGQK 234
Db 69 TSQRALEAPSGSSSTEWRNPNGNYGVTPTNTY-----RNSTGQYCREYTTQTVVIGGQK 123

QY 235 QEIYGTAC 242
Db 124 QKAYGNAC 131

RESULT 12
Q9F9Q9 PRELIMINARY; PRT; 151 AA.
AC Q9F9Q9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Outer membrane protein (fragment).
OS Rickettsia helvetica.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiaseae; Rickettsia.
OX NCBI_TaxID=35789;
RN [1]
RP SEQUENCE FROM N.A.
RA Nilsson K., Pahlson C.;
RT "Novel peptide diagnostic reagent and kit for detection of
RT rickettsiosis.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF181036; AAG09427.1; -.
FT NON TER 151
FT NON TER 151
SQ SEQUENCE 151 AA; 15621 MW; B77407B9C71E4B39 CRC64;

Query Match 17.9%; Score 239; DB 2; Length 151;
Best Local Similarity 35.9%; Pred. No. 3e-12;
Matches 55; Conservative 28; Mismatches 54; Indels 16; Gaps 6;

QY 103 SSLIITSV---PLVGC--AQNFSSRQEVGAATGAVVGVAGQLFGKSGRVSMATG-GAVL 156
Db 5 SKIMIIAALASMLQACNGGKNGKQGTGTLGGAGGALLGSQFGKGGQL-VGVGVGALL 63

QY 157 GGLIGSKIGQSMDOODK-----IKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYN 212
Db 64 GAVLGQIVAGHDEQDRRLAELTSQRALEAPSGSSSTEWRNPNGNYGVTPTNTY----- 119

QY 213 KQERQYQYCFQOKAMIAGQKQEIYGTACPOP 245
Db 120 -RNSTGQYCREYTTQTVVIGGQKQAYGNACROP 151

RESULT 13
Q9FQ01 PRELIMINARY; PRT; 131 AA.
AC Q9FQ01;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)

```

```

DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE 17 kDa protein (Fragment).
OS Rickettsia sp. California 2.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiaseae; Rickettsia.
OX NCBI_TaxID=147259;
RN [1]
RP SEQUENCE FROM N.A.
RA Raoult D.;
RC STRAIN=California 2;
RT "A new SFG rickettsia isolated from fleas.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Roux V., Raoult D.;
RC STRAIN=California 2;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF210693; AAG48554.1; -.
FT NON TER 1
FT NON TER 131
FT NON TER 131
SQ SEQUENCE 131 AA; 13374 MW; 23C8819B29FF860 CRC64;

Query Match 17.7%; Score 236.5; DB 2; Length 131;
Best Local Similarity 38.3%; Pred. No. 4e-12;
Matches 49; Conservative 24; Mismatches 44; Indels 11; Gaps 4;

QY 120 SRQEVGAATGAVVGVAGQLFGKSGRVSMATG-GAVLGGLIGSKIGQSMDOODK---I 174
Db 10 NKQGTGTLGGAGGALLGSQFGKGGQL-VGVGVGALLGVGGQIGAGMDEQDRRLAEL 68

QY 175 KLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKRQRRQYQYCFQOKAMIAGQK 234
Db 69 TSQRALEATPSGTSVEWRNPNGNHGYVTPNTY-----RNSTGQYCREYTTQTVVIGGQK 123

QY 235 QEIYGTAC 242
Db 124 QKAYGNAC 131

RESULT 14
Q52637 PRELIMINARY; PRT; 131 AA.
AC Q52637;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE 17 kDa antigen (Fragment).
OS Rickettsia sp.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiaseae; Rickettsia.
OX NCBI_TaxID=789;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=94117373; PubMed=8288533;
RA Werren J.H., Hurst G.D., Zhang W., Breeuwer J.A., Stouthamer R.,
RA Majerus M.E.;
RT "Rickettsial relative associated with male killing in the ladybird
RT beetle (Adalia bipunctata).";
RL J. Bacteriol. 176:388-394(1994).
DR EMBL; U04162; AAA19235.1; -.
FT NON TER 1
FT NON TER 131
FT NON TER 131
SQ SEQUENCE 131 AA; 13344 MW; AIDCF71050DF52DF CRC64;

Query Match 17.7%; Score 236.5; DB 2; Length 131;
Best Local Similarity 38.3%; Pred. No. 4e-12;
Matches 49; Conservative 24; Mismatches 44; Indels 11; Gaps 4;

QY 120 SRQEVGAATGAVVGVAGQLFGKSGRVSMATG-GAVLGGLIGSKIGQSMDOODK---I 174
Db 10 NKQGTGTLGGAGGALLGSQFGKGGQL-VGVGVGALLGVGGQIGAGMDEQDRRLAEL 68

QY 175 KLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKRQRRQYQYCFQOKAMIAGQK 234

```


GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 20:06:21 ; Search time 29.1969 Seconds
(without alignments)
875.264 Million cell updates/sec

Title: US-09-677-374-4

Perfect score: 832

Sequence: 1 MRGCLQGSSLIISVFLVGC.....IYGTACPODGRQWVISTEK 161

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_19Jun03:*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*

2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*

3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*

4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*

5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*

6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*

7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*

8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*

9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*

10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*

11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*

12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*

13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*

14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*

15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*

16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*

17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*

18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*

19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*

20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*

21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*

22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	832	100.0	161	22	AA881127
2	832	100.0	161	23	AAU97868
3	832	100.0	256	22	AA881128
4	832	100.0	256	23	AAU97869
5	815	98.0	162	22	AA878025
6	815	98.0	162	22	AA881126
7	815	98.0	162	23	AAU97867
8	127.5	15.3	182	24	ABJ18820
9	112	13.5	20	22	AA881130

10	112	13.5	20	23	AAU97871
11	111	13.3	197	23	AAO17565
12	111	13.3	224	22	AA820105
13	111	13.3	224	23	ABG80432
14	102.5	12.3	223	20	AAU34487
15	102.5	12.3	230	20	AAU34362
16	101	12.1	154	11	ABR05799
17	101	12.1	155	23	ABG80423
18	95	11.4	309	22	ABG15906
19	88	10.6	528	22	AB82611
20	87.5	10.5	147	22	AB83847
21	86.5	10.4	2309	22	AB866232
22	84	10.1	1251	22	AB861254
23	83	10.0	116	19	AAU11028
24	83	10.0	423	22	AA830695
25	82.5	9.9	71	24	ABP58756
26	82	9.9	112	23	ABP28121
27	82	9.9	2017	22	ABG08301
28	81.5	9.8	581	23	ABP25639
29	81	9.7	666	22	AB858019
30	80.5	9.7	269	23	AB890254
31	80.5	9.7	542	22	AB865790
32	80.5	9.7	542	22	AB865791
33	80.5	9.7	542	22	ABP70501
34	79.5	9.6	82	23	ABP28119
35	79.5	9.6	618	21	AA856803
36	79.5	9.6	1052	24	AAE33773
37	78.5	9.4	651	20	AAU40097
38	78.5	9.4	651	23	AAU11781
39	78.5	9.4	718	12	AA814308
40	78.5	9.4	718	19	AAW53346
41	78.5	9.4	718	21	AAU59070
42	78	9.4	240	24	ABP06080
43	77.5	9.3	102	22	AA841943
44	77.5	9.3	170	22	ABP37981
45	77.5	9.3	255	21	AA811398

ALIGNMENTS

RESULT 1

AA881127

ID AA881127 standard; Protein; 161 AA.

XX

AC AA881127;

XX

DT 11-JUL-2001 (first entry)

XX

DE Optimised OsPA protein 17E2 amino acid sequence.

XX

KW Poikilothermic fish; Piscirickettsia salmonis; rickettsial pathogen;

KW vaccine; OsPA; salmonid rickettsial septicaemia; rickettsial disease;

KW SRS.

XX

OS Piscirickettsia salmonis.

OS Synthetic.

XX

FH Key

FT Region

FT Location/Qualifiers

FT 109..128

FT /label= B_cell_epitope

XX

CA2281913-A1.

XX

PD 17-MAR-2001.

XX

PF C17E2 OsPA constru

XX

PR 17-SEP-1999; 99CA-2281913.

XX

PR 17-SEP-1999; 99CA-2281913.

XX

PA (KAYW/) KAY W W.

PA (BURI/) BURIAN J.

PA (KUZY/) KUZUK M A.

XX Kay WW, Burian J, Kuzyk MA;
 XX WPI; 2001-316844/34.
 XX N-PSDB; AAF86247.
 XX Method for protecting poikilothermic fish against salmonid rickettsial
 PT septicaemia and other rickettsial diseases comprises administering a
 PT vaccine containing the OspA protein of *Piscirickettsia salmonis* -
 XX
 XX Disclosure; Fig 5; 35pp; English.
 XX This invention relates to a method for the protection against infection
 CC of a poikilothermic fish by the bacterial pathogen, *Piscirickettsia*
 CC salmonis. The method comprises administering an immunogenic amount of a
 CC P. salmonis specific antigen termed OspA, or an immunogenic fragment of
 CC OspA in the form of a vaccine. The method is used for protecting animals,
 CC particularly poikilothermic fish, against the bacterial pathogen
 CC P. salmonis. The method is also useful for protecting against salmonid
 CC rickettsial septicaemia (SRS) and other rickettsial diseases. The present
 CC sequence represents optimised P. salmonis OspA protein 17E2. The DNA
 CC encoding OspA 17E2 (AAF86247) has been optimised for expression in
 CC *Escherichia coli*. An OspA protein with an N-terminal fusion partner is
 CC used in a vaccine to create an anti-OspA antibody response.
 XX
 XX Sequence 161 AA;
 SQ
 Query Match 100.0%; Score 832; DB 22; Length 161;
 Best Local Similarity 100.0%; Pred. No. 3.6e-81;
 Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRGCLQGSSLIISVFLVGCAGNFSRQEVGAATGAVVGGVAGQLFGKSGRVSMAIGGAV 60
 DB 1 MRGCLQGSSLIISVFLVGCAGNFSRQEVGAATGAVVGGVAGQLFGKSGRVSMAIGGAV 60
 QY 61 LGGILGSKIGSQMDQDKIKLNQSLKVKAGQVTRWRNPDGNSYVPEPRTYQRYNKOE 120
 DB 61 LGGILGSKIGSQMDQDKIKLNQSLKVKAGQVTRWRNPDGNSYVPEPRTYQRYNKOE 120
 QY 121 RROQYCREFOQKAMTAGQKQEIYGTACPDGRWQVISTEK 161
 DB 121 RROQYCREFOQKAMTAGQKQEIYGTACPDGRWQVISTEK 161
 RESULT 2
 AAU97868
 ID AAU97868 standard; Protein; 161 AA.
 AC AAU97868;
 XX 12-AUG-2002 (first entry)
 DT
 DE *Escherichia coli* codon optimised OspA, 17e2.
 XX Outer surface lipoprotein; OspA; antibacterial; immunosuppressive;
 KW vaccine; poikilothermic fish; fin-fish; Rickettsial septicaemia;
 KW Rickettsial disease; 17e2.
 XX
 OS *Piscirickettsia salmonis*.
 OS Synthetic.
 XX CA2339327-A1.
 XX 15-MAR-2002.
 XX 19-MAR-2001; 2001CA-2339327.
 XX 15-SEP-2000; 2000US-0677374.
 XX (THOR/) THORNTON J C.
 PA (KAYW/) KAY W W.
 PA (BURI/) BURIAN J.
 PA (KUZY/) KUZYK M A.

XX Thornton JC, Kay WW, Burian J, Kuzyk MA;
 XX WPI; 2002-455221/49.
 XX N-PSDB; ABK52402.
 XX Inducing immunity in fin fish to Rickettsial septicaemia, comprises
 PT administration of an outer surface lipoprotein (OspA) of a bacterial
 PT strain, as a vaccine -
 XX
 XX Example 4; Fig 5; 55pp; English.
 XX The invention describes a method of protecting a poikilothermic fish
 CC against infection by the bacterial pathogen *Piscirickettsia salmonis*
 CC comprising administering either intraperitoneally, by immersion or
 CC orally, an immunogenic amount of principal antigen, the OspA (outer
 CC surface lipoprotein), its variants, non-lipidated form or antigenic
 CC peptides derived or synthesized with or without an adjuvant. The new
 CC method is used to provide an outer surface lipoprotein (OspA) of
 CC bacterial strain *Piscirickettsia salmonis* as a vaccine to induce immunity
 CC in fin-fish against Rickettsial septicaemia and other related
 CC Rickettsial diseases caused by either a virus, bacteria or parasite.
 CC This is the amino acid sequence of the *Escherichia coli* codon optimised
 CC outer surface lipoprotein OspA (17e2) used in the creation of the vaccine
 CC described in the invention.
 XX
 XX Sequence 161 AA;
 SQ
 Query Match 100.0%; Score 832; DB 23; Length 161;
 Best Local Similarity 100.0%; Pred. No. 3.6e-81;
 Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRGCLQGSSLIISVFLVGCAGNFSRQEVGAATGAVVGGVAGQLFGKSGRVSMAIGGAV 60
 DB 1 MRGCLQGSSLIISVFLVGCAGNFSRQEVGAATGAVVGGVAGQLFGKSGRVSMAIGGAV 60
 QY 61 LGGILGSKIGSQMDQDKIKLNQSLKVKAGQVTRWRNPDGNSYVPEPRTYQRYNKOE 120
 DB 61 LGGILGSKIGSQMDQDKIKLNQSLKVKAGQVTRWRNPDGNSYVPEPRTYQRYNKOE 120
 QY 121 RROQYCREFOQKAMTAGQKQEIYGTACPDGRWQVISTEK 161
 DB 121 RROQYCREFOQKAMTAGQKQEIYGTACPDGRWQVISTEK 161
 RESULT 3
 AAB81128
 ID AAB81128 standard; Protein; 256 AA.
 XX AAB81128;
 AC AAB81128;
 XX 11-JUL-2001 (first entry)
 DT
 DE C17E2 OspA construct with N-terminal fusion partner.
 XX Poikilothermic fish; *Piscirickettsia salmonis*; rickettsial pathogen;
 KW vaccine; OspA; salmonid rickettsial septicaemia; rickettsial disease;
 KW SRS; 17E2; fusion construct.
 XX
 OS *Piscirickettsia salmonis*.
 OS Synthetic.
 XX Key Location/Qualifiers
 XX Region 1..95
 FT /label= Undefined_N-terminal_fusion_partner
 FT Region 96..256
 FT /label= C17E2_OspA
 FT /note= "Product of OspA gene optimised for expression in
 FT *Escherichia coli*"
 XX CA2281913-A1.
 XX 17-MAR-2001.
 PD

```

XX 17-SEP-1999; 99CA-2281913.
XX PF
XX PR
XX 17-SEP-1999; 99CA-2281913.
XX PA
XX (KAYW/) KAY W W.
XX (BURI/) BURIAN J.
XX (KUYZ/) KUYZK M A.
XX PI
XX Kay WW, Burian J, Kuzyk MA;
XX WPI; 2001-316844/34.
XX DR
XX N-PSDB; AAF86248.
XX PT
XX Method for protecting poikilothermic fish against salmonid rickettsial
XX PT
XX septicemia and other rickettsial diseases comprises administering a
XX PT
XX vaccine containing the OspA protein of Piscirickettsia salmonis -
XX XX
XX PS
XX Example 4; Fig 5; 35pp; English.
XX CC
XX This invention relates to a method for the protection against infection
XX CC
XX of a poikilothermic fish by the bacterial pathogen, Piscirickettsia
XX CC
XX salmonis. The method comprises administering an immunogenic amount of a
XX CC
XX P. salmonis specific antigen termed OspA, or an immunogenic fragment of
XX CC
XX OspA in the form of a vaccine. The method is used for protecting animals,
XX CC
XX particularly poikilothermic fish, against the bacterial pathogen
XX CC
XX P. salmonis. The method is also useful for protecting against salmonid
XX CC
XX rickettsial septicemia (SRS) and other rickettsial diseases. The present
XX CC
XX sequence represents the amino acid sequence of C17e2, a P. salmonis OspA
XX CC
XX construct optimised for expression in Escherichia coli, fused to an
XX CC
XX undefined N-terminal fusion partner. The fusion protein is used in a
XX CC
XX vaccine to create an anti-OspA antibody response.
XX XX
XX SQ
XX Sequence 256 AA;
XX
XX Query Match 100.0%; Score 832; DB 22; Length 256;
XX Best Local Similarity 100.0%; Pred. No. 6.6e-81;
XX Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 MRGCLQGSSLLIIISVFLVGCACQNFQSRQEVGAATGAVGVGAGQLFGKSGRVSMAIGGAV 60
XX Db 96 MRGCLQGSSLLIIISVFLVGCACQNFQSRQEVGAATGAVGVGAGQLFGKSGRVSMAIGGAV 155
XX
XX QY 61 LGGLIGSKIGQSMDDQDKIKLQNSLEKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKE 120
XX Db 156 LGGLIGSKIGQSMDDQDKIKLQNSLEKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKE 215
XX
XX QY 121 RRQYCREFOQKAMTAGQKQEIYGTACPPDGRWQVISTEK 161
XX Db 216 RRQYCREFOQKAMTAGQKQEIYGTACPPDGRWQVISTEK 256
XX
XX RESULT 4
XX AAU97869
XX ID AAU97869 standard; Protein; 256 AA.
XX AC
XX AAU97869;
XX XX
XX 12-AUG-2002 (first entry)
XX DT
XX E. coli codon optimised OspA, 17e2 with N-terminal fusion peptide.
XX DE
XX Outer surface lipoprotein; OspA; antibacterial; immunosuppressive;
XX KW
XX vaccine; poikilothermic fish; fin-fish; Rickettsial septicemia;
XX KW
XX Rickettsial disease; 17e2.
XX XX
XX Piscirickettsia salmonis.
XX OS
XX Synthetic.
XX XX
XX CA2339327-A1.
XX PN
XX 15-MAR-2002.
XX PD
XX
XX

```

```

PF 19-MAR-2001; 2001CA-2339327.
XX XX
XX 15-SEP-2000; 2000US-0677374.
XX XX
XX (THOR/) THORNTON J C.
XX PA
XX (KAYW/) KAY W W.
XX PA
XX (BURI/) BURIAN J.
XX PA
XX (KUYZ/) KUYZK M A.
XX XX
XX Thornton JC, Kay WW, Burian J, Kuzyk MA;
XX PI
XX WPI; 2002-455221/49.
XX DR
XX N-PSDB; ABK52403.
XX XX
XX Inducing immunity in fin fish to Rickettsial septicemia, comprises
XX PT
XX administration of an outer surface lipoprotein (OspA) of a bacterial
XX PT
XX strain, as a vaccine -
XX XX
XX PS
XX Example 4; Fig 5; 55pp; English.
XX CC
XX The invention describes a method of protecting a poikilothermic fish
XX CC
XX against infection by the bacterial pathogen Piscirickettsia Salmonis
XX CC
XX comprising administering either intraperitoneally, by immersion or
XX CC
XX orally, an immunogenic amount of principal antigen, the OspA (outer
XX CC
XX surface lipoprotein), its variants, non-lipidated form or antigenic
XX CC
XX peptides derived or synthesized with or without an adjuvant. The new
XX CC
XX method is used to provide an outer surface lipoprotein (OspA) of
XX CC
XX bacterial strain Piscirickettsia salmonis as a vaccine to induce immunity
XX CC
XX in fin-fish against Rickettsial septicemia and other related
XX CC
XX Rickettsial diseases caused by either a virus, bacteria or parasite.
XX CC
XX This is the amino acid sequence of the Escherichia coli codon optimised
XX CC
XX outer surface lipoprotein OspA (17e2) with an N-terminal fusion used in
XX CC
XX the creation of the vaccine described in the invention.
XX XX
XX SQ
XX Sequence 256 AA;
XX
XX Query Match 100.0%; Score 832; DB 23; Length 256;
XX Best Local Similarity 100.0%; Pred. No. 6.6e-81;
XX Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 MRGCLQGSSLLIIISVFLVGCACQNFQSRQEVGAATGAVGVGAGQLFGKSGRVSMAIGGAV 60
XX Db 96 MRGCLQGSSLLIIISVFLVGCACQNFQSRQEVGAATGAVGVGAGQLFGKSGRVSMAIGGAV 155
XX
XX QY 61 LGGLIGSKIGQSMDDQDKIKLQNSLEKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKE 120
XX Db 156 LGGLIGSKIGQSMDDQDKIKLQNSLEKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKE 215
XX
XX QY 121 RRQYCREFOQKAMTAGQKQEIYGTACPPDGRWQVISTEK 161
XX Db 216 RRQYCREFOQKAMTAGQKQEIYGTACPPDGRWQVISTEK 256
XX
XX RESULT 5
XX AAG78025
XX ID AAG78025 standard; Protein; 162 AA.
XX AC
XX AAG78025;
XX XX
XX 15-JAN-2002 (first entry)
XX DT
XX Piscirickettsia salmonis polypeptide P10.6.
XX DE
XX Piscirickettsia salmonis; piscirickettsiosis; salmonid rickettsial;
XX KW
XX septicemia; SRS; surface antigen; vaccine; antibacterial; fish;
XX KW
XX ATCC VR-1361.
XX XX
XX Piscirickettsia salmonis.
XX OS
XX WO200168865-A2.
XX PN
XX 20-SEP-2001.
XX PD
XX
XX

```

PF 12-MAR-2001; 2001WO-GB01055.
 XX
 PR 11-MAR-2000; 2000GB-0005838.
 PR 01-JUL-2000; 2000GB-0016080.
 PR 01-JUL-2000; 2000GB-0016082.
 PR 29-JUL-2000; 2000GB-0018599.
 XX
 PA (AQUA-) AQUA HEALTH EURO LTD.
 XX
 XX Simard N, Brouwers H, Jones S, Griffiths S, Valenzuela P;
 PI Burzio L;
 XX WPI; 2001-639050/73.
 DR N-PSDB; AAH79040.
 XX
 XX New nucleic acids encoding an amino acid sequence homologous to the
 PT surface antigen present on Piscirickettsia salmonis are useful to
 PT protect fish against piscirickettsiosis -
 XX
 XX Claim 6; Fig 5; 25pp; English.
 XX
 CC The invention relates to nucleic acid sequences and the encoded protein
 CC of a least part of the surface antigen present on Piscirickettsia
 CC salmonis for production of a vaccine with antibacterial activity to
 CC protect fish against P. salmonis which causes piscirickettsiosis, also
 CC known as salmonid rickettsial septicaemia.
 XX
 XX Sequence 162 AA;
 SQ
 Query Match 98.0%; Score 815; DB 22; Length 162;
 Best Local Similarity 98.8%; Pred. No. 2.4e-79;
 Matches 158; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 2 RGCLOGSSLLIIISVFLVGCQNFSRQEVGAATGAVVGGVAGQLFGKSGRVSMAIGGAVL 61
 DB 3 RGCLOGSSLLIIISVFLVGCQNFSRQEVGAATGAVVGGVAGQLFGKSGRVSMAIGGAVL 62
 QY 62 GGLIGSKIGQSMDDQDKIKLQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKOER 121
 DB 63 GGLIGSKIGQSMDDQDKIKLQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKOER 122
 QY 122 RQCYCREFOQKAMIAQKQEIYGTACPDGRWQVISTEK 161
 DB 123 RQCYCREFOQKAMIAQKQEIYGTACPDGRWQVISTEK 162
 RESULT 6
 AAB81126
 ID AAB81126 standard; Protein; 162 AA.
 XX
 AC AAB81126;
 XX
 DT 11-JUL-2001 (first entry)
 XX
 DE OSPA antigen amino acid sequence.
 XX
 XX Poikilothermic fish; Piscirickettsia salmonis; rickettsial pathogen;
 KW vaccine; OSPA; salmonid rickettsial septicaemia; rickettsial disease;
 KW SR.
 XX
 XX Piscirickettsia salmonis.
 OS
 XX
 XX Key Location/Qualifiers
 FT Region 110..129
 FT /label= B_cell_epitope
 FT
 XX CA2281913-A1.
 XX
 XX 17-MAR-2001.
 PD
 XX 17-SEP-1999; 99CA-2281913.
 PF
 XX 17-SEP-1999; 99CA-2281913.
 PR

XX (KAYW/) KAY W W.
 PA (BURI/) BURIAN J.
 PA (KUYZ/) KUYZ M A.
 XX
 XX Kay WW, Burian J, Kuzyk MA;
 PI
 XX WPI; 2001-316844/34.
 DR N-PSDB; AAF86246.
 XX
 XX Method for protecting poikilothermic fish against salmonid rickettsial
 PT septicaemia and other rickettsial diseases comprises administering a
 PT vaccine containing the OSPA protein of Piscirickettsia salmonis -
 XX
 XX Example 2; Fig 2B; 35pp; English.
 XX
 CC This invention relates to a method for the protection against infection
 CC of a poikilothermic fish by the bacterial pathogen, Piscirickettsia
 CC salmonis. The method comprises administering an immunogenic amount of a
 CC P. salmonis specific antigen termed OSPA, or an immunogenic fragment of
 CC OSPA in the form of a vaccine. The method is used for protecting animals,
 CC particularly poikilothermic fish, against the bacterial pathogen
 CC P. salmonis. The method is also useful for protecting against salmonid
 CC rickettsial septicaemia (SR) and other rickettsial diseases. The present
 CC sequence represents P. salmonis OSPA protein. An OSPA protein with an
 CC N-terminal fusion partner is used in a vaccine to create an anti-OSPA
 CC antibody response.
 XX
 XX Sequence 162 AA;
 SQ
 Query Match 98.0%; Score 815; DB 22; Length 162;
 Best Local Similarity 98.8%; Pred. No. 2.4e-79;
 Matches 158; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 2 RGCLOGSSLLIIISVFLVGCQNFSRQEVGAATGAVVGGVAGQLFGKSGRVSMAIGGAVL 61
 DB 3 RGCLOGSSLLIIISVFLVGCQNFSRQEVGAATGAVVGGVAGQLFGKSGRVSMAIGGAVL 62
 QY 62 GGLIGSKIGQSMDDQDKIKLQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKOER 121
 DB 63 GGLIGSKIGQSMDDQDKIKLQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKOER 122
 QY 122 RQCYCREFOQKAMIAQKQEIYGTACPDGRWQVISTEK 161
 DB 123 RQCYCREFOQKAMIAQKQEIYGTACPDGRWQVISTEK 162
 RESULT 7
 AAU97867
 ID AAU97867 standard; Protein; 162 AA.
 XX
 AC AAU97867;
 XX
 DT 12-AUG-2002 (first entry)
 XX
 DE Piscirickettsia salmonis outer surface lipoprotein OSPA.
 XX
 XX Outer surface lipoprotein; OSPA; antibacterial; immunosuppressive;
 KW vaccine; poikilothermic fish; fin-fish; Rickettsial septicaemia;
 KW Rickettsial disease.
 XX
 XX Piscirickettsia salmonis.
 OS
 XX
 XX CA2339327-A1.
 XX
 XX 15-MAR-2002.
 PD
 XX 19-MAR-2001; 2001CA-2339327.
 PF
 XX 15-SEP-2000; 2000US-0677374.
 PR
 XX (THOR/) THORNTON J C.
 XX (KAYW/) KAY W W.
 PA

PA (BURI/) BURIAN J.
 XX (KUZY/) KUYZK M A.
 XX Thornton JC, Kay WW, Burian J, Kuzyk MA;
 XX WPI; 2002-455221/49.
 DR N-PSDB; ABK52401.
 XX Inducing immunity in fin fish to Rickettsial septicaemia, comprises
 PT administration of an outer surface lipoprotein (OspA) of a bacterial
 PT strain, as a vaccine -
 XX
 PS Claim 15; Fig 2; 55pp; English.
 XX
 CC The invention describes a method of protecting a poikilothermic fish
 CC against infection by the bacterial pathogen *Piscirickettsia salmonis*
 CC comprising administering either intraperitoneally, by immersion or
 CC orally, an immunogenic amount of principal antigen, the OspA (outer
 CC surface lipoprotein), its variants, non-lipidated form or antigenic
 CC peptides derived or synthesized with or without an adjuvant. The new
 CC method is used to provide an outer surface lipoprotein (OspA) of
 CC bacterial strain *Piscirickettsia salmonis* as a vaccine to induce immunity
 CC in fin-fish against Rickettsial septicaemia and other related
 CC Rickettsial diseases caused by either a virus, bacteria or parasite.
 CC This is the amino acid sequence of the *Piscirickettsia salmonis* outer
 CC surface lipoprotein, OspA, used in the creation of the vaccine described
 CC in the invention.
 XX
 SQ Sequence 162 AA;
 Query Match 98.0%; Score 815; DB 23; Length 162;
 Best Local Similarity 98.8%; Pred. No. 2.4e-79;
 Matches 158; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 2 RGLQSSLIISVFLVGCQAQNFROEVGAATGAVVGGVAGQLFGKSGRVSMAIGGAVL 61
 DB 3 RGLQSSLIISVFLVGCQAQNFROEVGAATGAVVGGVAGQLFGKSGRVSMAIGGAVL 62
 QY 62 GGLIGSKIGOSMDQDDKIKLNOSLEKVKAGQVTRWRNPDTGNSYVEPVRTYQYKNQER 121
 DB 63 GGLIGSKIGOSMDQDDKIKLNOSLEKVKAGQVTRWRNPDTGNSYVEPVRTYQYKNQER 122
 QY 122 RQCYREFQOKAMIAQKQEIYGTACPDGRWQVISTEK 161
 DB 123 RQCYREFQOKAMIAQKQEIYGTACPDGRWQVISTEK 162
 RESULT 8
 ABJ18820
 ID ABJ18820 standard; Protein; 182 AA.
 XX
 AC ABJ18820;
 XX
 DT 27-FEB-2003 (first entry)
 XX
 DE Pseudomonas aeruginosa biofilm formation-related protein #84.
 XX
 KW Biofilm formation modulation; biofilm-associated disease;
 KW Cystic fibrosis; AIDS; middle ear infection; acne; periodontal disease;
 KW catheter-associated infection; medical device-associated infection.
 XX
 OS Pseudomonas aeruginosa.
 XX
 PN WO200285295-A2.
 XX
 PD 31-OCT-2002.
 XX
 PF 19-APR-2002; 2002WO-US12532.
 XX
 PR 20-APR-2001; 2001US-285190P.
 PR 24-OCT-2001; 2001US-344142P.
 XX
 PA (IOWA) UNIV IOWA RES FOUND.

PA (HARD) HARVARD COLLEGE.
 XX
 PI Whiteley M, Bangera MG, Lory S, Greenberg EP;
 XX
 DR WPI; 2003-075601/07.
 DR N-PSDB; ABT14642.
 XX
 PT Identifying compound capable of modulating biofilm formation by
 PT bacteria/bacterial antibiotic resistance, useful for treatment of
 PT biofilm associated disease -
 XX
 PS Claim 4; Page 152; 154pp; English.
 XX
 CC The invention comprises a method for identifying a compound capable of
 CC modulating biofilm formation by bacteria. The method of the invention is
 CC useful for identifying a compound capable of modulating biofilm formation
 CC by bacteria or modulating bacterial antibiotic resistance. The method of
 CC the invention is also useful for diagnosing and treating a subject
 CC (especially an immunocompromised human) that is afflicted with a biofilm-
 CC associated disease or disorder, such as: cystic fibrosis; AIDS; middle
 CC ear infections; acne; periodontal disease; catheter-associated
 CC infections; and medical device-associated infections. The present amino
 CC acid sequence represents a protein that is used in the invention.
 XX
 SQ Sequence 182 AA;
 Query Match 15.3%; Score 127.5; DB 24; Length 182;
 Best Local Similarity 34.8%; Pred. No. 1.8e-05;
 Matches 32; Conservative 13; Mismatches 28; Indels 19; Gaps 2;
 QY 27 QEVGAATGAVVGGVAGQLFGKSGRVSMAIGGAVLGLIGSKIGOSMDQDDKI----- 79
 DB 70 QIAGTAIGAAGVGGLLGNQIGGTGKKIATVAGAGVGGVAGNVQEGMQERDVTYTTETRC 129
 QY 80 -KLNOSLEKV-----KAGQVTRWRNP 99
 DB 130 STVHDSSEKVVGYDVKYMLDGGKAGQIRMERDP 161
 RESULT 9
 AB881130
 ID AAB81130 standard; Peptide; 20 AA.
 XX
 AC AAB81130;
 XX
 DT 11-JUL-2001 (first entry)
 XX
 DE OspA B-cell epitope peptide #2.
 XX
 KW Poikilothermic fish; *Piscirickettsia salmonis*; rickettsial pathogen;
 KW vaccine; OspA; salmonid rickettsial septicaemia; rickettsial disease;
 KW SRS; antibody.
 XX
 OS *Piscirickettsia salmonis*.
 XX
 PN CA2281913-A1.
 XX
 PD 17-MAR-2001.
 XX
 PF 17-SEP-1999; 99CA-2281913.
 XX
 PR 17-SEP-1999; 99CA-2281913.
 XX
 PA (KAYW/) KAY W W.
 PA (BURI/) BURIAN J.
 PA (KUZY/) KUYZK M A.
 XX
 PI Kay WW, Burian J, Kuzyk MA;
 XX
 DR WPI; 2001-316844/34.
 XX
 PT Method for protecting poikilothermic fish against salmonid rickettsial
 PT septicaemia and other rickettsial diseases comprises administering a

PT vaccine containing the OspA protein of Piscirickettsia salmonis -
 PS Example 2; Page 17; 35pp; English.
 XX
 CC This invention relates to a method for the protection against infection
 CC of a poikilothermic fish by the bacterial pathogen, Piscirickettsia
 CC salmonis. The method comprises administering an immunogenic amount of a
 CC P. salmonis specific antigen termed OspA, or an immunogenic fragment of a
 CC OspA in the form of a vaccine. The method is used for protecting animals,
 CC particularly poikilothermic fish, against the bacterial pathogen
 CC P. salmonis. The method is also useful for protecting against salmonid
 CC rickettsial septicemia (SRs) and other rickettsial diseases. The present
 CC sequence represents an immunogenic epitope of the P. salmonis OspA
 CC protein. The peptide is used to raise rabbit anti-OspA antibodies.
 XX
 SQ Sequence 20 AA;
 Query Match 13.5%; Score 112; DB 22; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.7e-05;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 109 PVRTYQRYNKQERRQOYCCE 128
 DB 1 PVRTYQRYNKQERRQOYCCE 20
 RESULT 10
 AAU97871
 ID AAU97871 standard; Peptide; 20 AA.
 AC AAU97871;
 DT 12-AUG-2002 (first entry)
 DE Outer surface lipoprotein OspA based peptide #2.
 XX
 KW Outer surface lipoprotein; OspA; antibacterial; immunosuppressive;
 KW vaccine; poikilothermic fish; fin-fish; Rickettsial septicemia;
 KW Rickettsial disease.
 XX
 OS Piscirickettsia salmonis.
 OS Synthetic.
 XX
 CA2339327-A1.
 XX
 PD 15-MAR-2002.
 XX
 PF 19-MAR-2001; 2001CA-2339327.
 XX
 PR 15-SEP-2000; 2000US-0677374.
 XX
 PA (THOR/) THORNTON J C.
 PA (KAYW/) KAY W W.
 PA (BURI/) BURIAN J.
 PA (KUZY/) KUZUK M A.
 XX
 PI Thornton JC, Kay WW, Burian J, Kuzyk MA;
 PI WPI; 2002-455221/49.
 XX
 DR Inducing immunity in fin fish to Rickettsial septicemia, comprises
 PT administration of an outer surface lipoprotein (OspA) of a bacterial
 PT strain, as a vaccine -
 XX
 PS Example 2; Page 17; 55pp; English.
 CC
 CC The invention describes a method of protecting a poikilothermic fish
 CC against infection by the bacterial pathogen Piscirickettsia Salmonis
 CC comprising administering either intraperitoneally, by immersion or
 CC orally, an immunogenic amount of principal antigen, the OspA (outer
 CC surface lipoprotein), its variants, non-lipidated form or antigenic
 CC peptides derived or synthesized with or without an adjuvant. The new
 CC method is used to provide an outer surface lipoprotein (OspA) of

CC bacterial strain Piscirickettsia salmonis as a vaccine to induce immunity
 CC in fin-fish against Rickettsial septicemia and other related
 CC Rickettsial diseases caused by either a virus, bacteria or parasite.
 CC This sequence represents a synthetic peptide used to create polyclonal
 CC rabbit antibodies against the Piscirickettsia salmonis outer surface
 CC lipoprotein, OspA.
 XX
 SQ Sequence 20 AA;
 Query Match 13.5%; Score 112; DB 23; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.7e-05;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 109 PVRTYQRYNKQERRQOYCCE 128
 DB 1 PVRTYQRYNKQERRQOYCCE 20
 RESULT 11
 AAO17565
 ID AAO17565 standard; Protein; 197 AA.
 XX
 AC AAO17565;
 XX
 DT 19-JUL-2002 (first entry)
 XX
 DE M catarrhalis MCAL00414 protein SEQ ID NO: 10.
 XX
 KW Moraxella; vaccine; respiratory tract infection; antiinflammatory;
 KW auditory; antibacterial; otitis media; sinusitis; pneumonia.
 XX
 OS Moraxella catarrhalis.
 XX
 PN WO200218595-A2.
 XX
 PD 07-MAR-2002.
 XX
 PF 28-AUG-2001; 2001WO-CA01221.
 XX
 PR 28-AUG-2000; 2000US-228294P.
 PR 28-AUG-2000; 2000US-228295P.
 PR 28-AUG-2000; 2000US-228296P.
 PR 29-AUG-2000; 2000US-228438P.
 PR 29-AUG-2000; 2000US-228439P.
 PR 29-AUG-2000; 2000US-228440P.
 PR 29-AUG-2000; 2000US-228441P.
 PR 29-AUG-2000; 2000US-228442P.
 PR 29-AUG-2000; 2000US-228443P.
 PR 29-AUG-2000; 2000US-228511P.
 PR 29-AUG-2000; 2000US-228512P.
 PR 29-AUG-2000; 2000US-228742P.
 PR 01-SEP-2000; 2000US-228773P.
 PR 01-SEP-2000; 2000US-229465P.
 PR 01-SEP-2000; 2000US-229474P.
 PR 01-SEP-2000; 2000US-229475P.
 PR 05-SEP-2000; 2000US-229740P.
 PR 05-SEP-2000; 2000US-229803P.
 PR 05-SEP-2000; 2000US-229804P.
 PR 05-SEP-2000; 2000US-229805P.
 PR 05-SEP-2000; 2000US-229806P.
 PR 05-SEP-2000; 2000US-229809P.
 PR 06-SEP-2000; 2000US-229811P.
 PR 06-SEP-2000; 2000US-230214P.
 PR 06-SEP-2000; 2000US-230250P.
 PR 06-SEP-2000; 2000US-230252P.
 XX
 PA (AVET) AVENTIS PASTEUR LTD.
 XX
 PI Loosmore S, Wang J, Bradley B, Ochs M, Yang Y;
 XX WPI; 2002-401721/43.
 DR N-PSDB; AAL46497.

CC readily filter sterilised. The blebs can be made and harvested without
 CC the use of detergents such as deoxycholate, thus obviating chromatography
 CC purification and ultra centrifugation steps. Vesicles prepared from the
 CC invention have reduced particle size (allowing sterile filtration through
 CC 0.22 µm pores), increased batch homogeneity, and a superior yield.
 CC The present amino acid sequence represents a Moraxella catarrhalis
 CC protein, as described in the invention.

XX
 SQ Sequence 224 AA;

Query Match 13.3%; Score 111; DB 23; Length 224;
 Best Local Similarity 33.3%; Pred. No. 0.0014;
 Matches 37; Conservative 15; Mismatches 35; Indels 24; Gaps 5;

QY 7 GSSLIIISVFLVGCQAQNFSSRQEVCAATG-----AVVGGVAGQLFGKSGRVSMIAIGG 58

Db 7 GVVLASSMALAGCANTGT---TNGTGFGANVNVKAVIGAVAGAL---GGTAISKATGG 60

QY 59 -----AVLGLIGSKIGQSMDDQDKIKLNQSLKVKAGQVTRWRNPDTGN 103

Db 61 EKTGRDAILGAAGVAAAGAYMERQAK----QIEQQMQGTGTVTHTDITGN 107

RESULT 14

AAV34487
 ID AAY34487 standard; Protein; 223 AA.

XX
 AC AAY34487;

XX
 DT 20-MAR-2003 (updated)

DT 25-AUG-1999 (first entry)

XX
 XX

DE Porphorymonas gingivalis protein PG3.

XX
 KW Porphorymonas gingivalis; PG; periodontal disease; gingivitis;
 KW vaccine; antigenic.

XX
 OS Porphorymonas gingivalis.

XX
 PN WO9929870-A1.

XX
 PD 17-JUN-1999.

XX
 PF 10-DEC-1998; 98WO-AU01023.

XX
 PR 10-DEC-1997; 97AU-0000839.

PR 31-DEC-1997; 97AU-0001182.

PR 30-JAN-1998; 98AU-0001346.

PR 10-MAR-1998; 98AU-0002264.

PR 09-APR-1998; 98AU-0002911.

PR 23-APR-1998; 98AU-0003128.

PR 05-MAY-1998; 98AU-0003338.

PR 22-MAY-1998; 98AU-0003654.

PR 29-JUL-1998; 98AU-0004917.

PR 30-JUL-1998; 98AU-0004963.

PR 04-AUG-1998; 98AU-0005028.

XX
 PA (CSLC-) CSL LTD.

XX
 PI Agius CT, Barr IG, Hocking DM, Margetts MB, Patterson MA;

PI Ross BC, Rothel LJ, Webb EA;

XX
 WPI; 1999-385613/32.

DR N-PSDB; AAX91705.

XX
 PT Antigenic Porphorymonas gingivalis peptides for preventing
 PT gingivitis

XX
 PS Claim 1; Page 469; 588pp; English.

XX
 CC AAX91536 to AAX91801 encode two hundred and sixty six antigenic

CC Porphorymonas gingivalis (PG) polypeptide sequences given in AAY34318 to

CC AAY34583. AAX91802 to AAX91989 represent PCR primers used in the

CC isolation of the PG polypeptides. The PG polypeptides have antibacterial
 CC activity with a vaccine mechanism of action. The PG polypeptides can be
 CC used as vaccines especially against Porphorymonas gingivalis. Probes can
 CC be used to detect Porphorymonas gingivalis in standard hybridisation
 CC assays. Porphorymonas gingivalis is involved in periodontal disease
 CC especially gingivitis.

XX
 CC (Updated on 20-MAR-2003 to correct PR field.)

SQ Sequence 223 AA;

Query Match 12.3%; Score 102.5; DB 20; Length 223;
 Best Local Similarity 34.7%; Pred. No. 0.011;
 Matches 25; Conservative 13; Mismatches 31; Indels 3; Gaps 2;

QY 8 SSSLIIISVFLVGCQAQNFSSRQEVCAATGAVVGGVAGQLFGKSGRVSM-AIGCAVLGLIG 66

Db 8 ASVLVALVFAFCGLN--NMAKGLIGAGVGAIGAGVGNVAGNTAVGATVAVGGAAG 65

QY 67 SKIQSMDQDDK 78

Db 66 ALIGKKMDKQK 77

RESULT 15

AAV34362

ID AAY34362 standard; Protein; 230 AA.

XX
 AC AAY34362;

XX
 DT 20-MAR-2003 (updated)

DT 25-AUG-1999 (first entry)

XX
 XX

DE Porphorymonas gingivalis protein PG3.

XX
 KW Porphorymonas gingivalis; PG; periodontal disease; gingivitis;
 KW vaccine; antigenic.

XX
 OS Porphorymonas gingivalis.

XX
 PN WO9929870-A1.

XX
 PD 17-JUN-1999.

XX
 PF 10-DEC-1998; 98WO-AU01023.

XX
 PR 10-DEC-1997; 97AU-0000839.

PR 31-DEC-1997; 97AU-0001182.

PR 30-JAN-1998; 98AU-0001546.

PR 10-MAR-1998; 98AU-0002264.

PR 09-APR-1998; 98AU-0002911.

PR 23-APR-1998; 98AU-0003128.

PR 05-MAY-1998; 98AU-0003338.

PR 22-MAY-1998; 98AU-0003654.

PR 29-JUL-1998; 98AU-0004917.

PR 30-JUL-1998; 98AU-0004963.

PR 04-AUG-1998; 98AU-0005028.

XX
 PA (CSLC-) CSL LTD.

XX
 PI Agius CT, Barr IG, Hocking DM, Margetts MB, Patterson MA;

PI Ross BC, Rothel LJ, Webb EA;

XX
 WPI; 1999-385613/32.

DR N-PSDB; AAX91580.

XX
 PT Antigenic Porphorymonas gingivalis peptides for preventing
 PT gingivitis

XX
 PS Claim 1; Page 325-326; 588pp; English.

XX
 CC AAX91536 to AAX91801 encode two hundred and sixty six antigenic

CC Porphorymonas gingivalis (PG) polypeptide sequences given in AAY34318 to

CC AAY34583. AAX91802 to AAX91989 represent PCR primers used in the

CC isolation of the PG polypeptides. The PG polypeptides have antibacterial
CC activity with a vaccine mechanism of action. The PG polypeptides can be
CC used as vaccines especially against Porphyromonas gingivalis. Probes can
CC be used to detect Porphyromonas gingivalis in standard hybridisation
CC assays. Porphyromonas gingivalis is involved in periodontal disease
CC especially gingivitis.
CC (Updated on 20-MAR-2003 to correct PR field.)
XX

SQ Sequence 230 AA;
Query Match 12.3%; Score 102.5; DB 20; Length 230;
Best Local Similarity 34.7%; Pred. No. 0.012;
Matches 25; Conservative 13; Mismatches 31; Indels 3; Gaps 2;
QY 8 SLLIIISVFLVGCQNFQRQEVGAATGAVVGGVAGQVGFQKSGRVSM-AIGGAVLGGGLIG 66
Db 15 ASVLAVALVFAGGCLN--NMAKGLLIGAGVGGAGIAGVGNVAGNTAVGAIVGTAVGGAG 72
QY 67 SKIGSMDQDDK 78
Db 73 ALIGKMDKQKK 84

Search completed: November 5, 2003, 20:12:50
Job time : 30.1969 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 20:11:02 ; Search time 10.2884 Seconds
(without alignments)
662.108 Million cell updates/sec

Title: US-09-677-374-4

Perfect score: 832

Sequence: 1 MRGCLQGSSLIISVFLVGC.....IYGACPDGRWQVISTEK 161

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:*
- 1: /cgn2_6/ptodata/1/1aa/5A-COMB.pep.*
 - 2: /cgn2_6/ptodata/1/1aa/5B-COMB.pep.*
 - 3: /cgn2_6/ptodata/1/1aa/6A-COMB.pep.*
 - 4: /cgn2_6/ptodata/1/1aa/6B-COMB.pep.*
 - 5: /cgn2_6/ptodata/1/1aa/PCTUS-COMB.pep.*
 - 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	127.5	15.3	192	4	US-09-252-991A-22817
2	111.5	13.4	165	4	US-09-252-991A-19701
3	99.5	12.0	306	4	US-09-252-991A-21754
4	99.5	12.0	387	4	US-09-252-991A-23334
5	92.5	11.1	112	4	US-09-252-991A-31960
6	92.5	11.1	199	4	US-09-328-352-7802
7	90	10.8	126	4	US-09-328-352-8009
8	86.5	10.4	127	4	US-09-252-991A-28397
9	86.5	10.4	730	4	US-09-328-352-4442
10	86	10.3	1034	4	US-09-252-991A-26658
11	85	10.2	258	4	US-09-328-352-4253
12	83.5	10.0	84	4	US-09-107-532A-5149
13	82.5	9.9	150	4	US-09-328-352-6423
14	82.5	9.9	217	4	US-09-328-352-7068
15	82	9.9	215	4	US-09-328-352-6750
16	82	9.9	387	4	US-09-328-352-6442
17	82	9.9	734	4	US-09-252-991A-33036
18	81.5	9.8	104	4	US-09-107-532A-7222
19	80	9.6	572	4	US-09-252-991A-23878
20	79.5	9.6	259	4	US-09-328-352-5775
21	79	9.5	172	4	US-09-252-991A-27926
22	78.5	9.4	82	4	US-09-107-532A-4748
23	78.5	9.4	651	3	US-08-556-978B-19
24	78.5	9.4	651	3	US-09-247-806-1
25	78.5	9.4	718	1	US-08-425-069-2
26	78.5	9.4	718	2	US-08-317-844B-2
27	78.5	9.4	747	3	US-09-034-177-3

28	77.5	9.3	255	4	US-09-553-498-8	Sequence 8, Appli
29	77.5	9.3	255	4	US-09-618-869-8	Sequence 8, Appli
30	77.5	9.3	518	4	US-09-252-991A-19162	Sequence 19162, A
31	77.5	9.3	849	4	US-09-252-991A-31525	Sequence 31525, A
32	77.5	9.3	1415	4	US-09-252-991A-26438	Sequence 26438, A
33	77	9.3	397	4	US-09-252-991A-30059	Sequence 30059, A
34	77	9.3	551	2	US-09-067-351-2	Sequence 2, Appli
35	77	9.3	551	3	US-09-360-490-2	Sequence 2, Appli
36	76.5	9.2	2516	3	US-08-374-077C-2	Sequence 2, Appli
37	76.5	9.2	2516	3	US-08-895-590-2	Sequence 2, Appli
38	76.5	9.2	2516	4	US-09-539-879A-2	Sequence 2, Appli
39	76	9.1	141	4	US-09-328-352-7100	Sequence 7100, Ap
40	76	9.1	272	4	US-09-252-991A-29681	Sequence 29681, A
41	76	9.1	865	1	US-07-803-633A-13	Sequence 13, Appl
42	75.5	9.1	204	4	US-09-107-532A-6584	Sequence 6584, Ap
43	75	9.0	141	4	US-09-252-991A-23427	Sequence 23427, A
44	75	9.0	414	4	US-09-252-991A-27975	Sequence 27975, A
45	75	9.0	465	4	US-09-252-991A-18919	Sequence 18919, A

RESULT 1

US-09-252-991A-22817

; Sequence 22817, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 22817

; LENGTH: 192

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-22817

Query Match 15.3%; Score 127.5; DB 4; Length 192;

Best Local Similarity 34.8%; Pred. No. 1.4e-06;

Matches 32; Conservative 13; Mismatches 28; Indels 19; Gaps 2;

QY 27 QEVGAATGAVGVGGVAGQLFGKSGRVSMAIGAVLGLIGSKIGOSMDQDKI-----79

DB 80 QIATGTAIGAVVGLLGNQIGGTGKTIATVAGVGGVAGNVKVGQMDRTYTTTETRC 139

QY 80 -KLNSLEKV-----KAGQVTRWRNP 99

DB 140 STVHDSSEKVVGYDVVKYMLDGKAGQIRMERDP 171

RESULT 2

US-09-252-991A-19701

; Sequence 19701, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27


```

; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26658

Query Match      10.3%; Score 86; DB 4; Length 1034;
Best Local Similarity 34.8%; Pred. No. 0.64;
Matches 24; Conservative 10; Mismatches 29; Indels 6; Gaps 2;

QY 21 ANFGRQEVGAATGAVGVAGQVLFK-----GSRVSMAGAVLGGVIGLIGSKIGQSMQDQ 75
Db 866 AHAIKRGDVGGSFNSLVGLVGHNGGELVNVVDASGRVSAASASV-GGLVGSNAGSILSA 924

QY 76 QDKIKLNOS 84
Db 925 RSSSTVNGS 933

RESULT 11
US-09-328-352-4253
; Sequence 4253, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4253
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4253

Query Match      10.2%; Score 85; DB 4; Length 258;
Best Local Similarity 29.7%; Pred. No. 0.13;
Matches 22; Conservative 14; Mismatches 24; Indels 14; Gaps 2;

QY 29 VGAATGAVGVAGQVLFK-----GSRVSMAGAVLGGVIGLIGSKIGQSMQDQ 75
Db 166 VSPATAGTVGVTGAIIGKFNSSGNMSYQATGAGAGGAIIGLIVAAITNAEVGKIIG 225

QY 76 QDKIKLNOSLEKVK 89
Db 226 L-PIKESSEFMELK 238

RESULT 12
US-09-107-532A-5149
; Sequence 5149, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS: 7310
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598

```

```

; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5149:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...84
; SEQUENCE DESCRIPTION: SEQ ID NO: 5149:
US-09-107-532A-5149

Query Match      10.0%; Score 83.5; DB 4; Length 84;
Best Local Similarity 42.5%; Pred. No. 0.043;
Matches 17; Conservative 9; Mismatches 11; Indels 3; Gaps 1;

QY 34 GAVVGVAGQVLFKSGRVSMAIGGAVLGGVIGLIGSKIGQSM 73
Db 14 GGIIGATAGAITNRGS---SMGIITANVAVGLVGSALQAL 50

RESULT 13
US-09-328-352-6423
; Sequence 6423, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6423
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6423

Query Match      9.9%; Score 82.5; DB 4; Length 150;
Best Local Similarity 37.7%; Pred. No. 0.12;
Matches 20; Conservative 6; Mismatches 20; Indels 7; Gaps 1;

QY 30 GAATGAVVGVAGQVLFKSGRVSMAIGGAVLGGVIGLIGSKIGQSMQDQ 75
Db 79 GATVGAIGTVAGPLGVIGTVGVTFVGAISAGITGGVVGIVFGKAGVWIDK 131

RESULT 14
US-09-328-352-7068
; Sequence 7068, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252

```



```
; SEQ ID NO 7068
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7068

Query Match          9.9%; Score 82.5; DB 4; Length 217;
Best Local Similarity 25.2%; Pred. No. 0.2;
Matches 36; Conservative 20; Mismatches 52; Indels 35; Gaps 4;

QY 5 LQSSLLIIISVFLVGCAGNFSRQEVGAATGAVVGGVAGQLFGKSG-----50
Db 57 LASALIATTSMVTVAHADNGTRVAATSAIGSVVGTGKSGIGTSGATIGALGAGGAA 116
QY 51 -----RVSMAGGAVLGGIGSKIGSQMDQDKIKLNSLEKVKAG----QVTRWRNP 99
Db 117 AASDRNTEAAIGGA-LGGGAGYTVGNMGGTNGGYICAAVGAAGGSGALGRKVAQDRNY 175
QY 100 DTGNSYSVEPVRTYORYNKQERR 122
Db 176 DD-----RYDRRYDRDRR 189

RESULT 15
US-09-328-352-6750
; Sequence 6750, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6750
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6750

Query Match          9.9%; Score 82; DB 4; Length 215;
Best Local Similarity 23.9%; Pred. No. 0.22;
Matches 34; Conservative 22; Mismatches 48; Indels 38; Gaps 7;

QY 7 GSSLIIISVFLVGCAGNFSRQEVGAATGAVVGGVAGQLFGKSGRVSMAGGAVLGGIG 66
Db 68 GENLVATGAGTLGGA-----AVGAAGF--WGGPPGAVVG-----GIIGGVVGAIA 112
QY 67 SKIGSQMDQDKIKLNSLEKVKAGQVTRWRNPDT--GNSYSVEPVRTYORYNKQERRQQ 124
Db 113 NDIAGTNNQKD-----DSNDWQEDNYRWENYKMP-----YSEDKNLE 152
QY 125 YCREFOQKAMTAGQKOEIYGTA 146
Db 153 YDRDY-RAAYRLGYENRVHNA 173
```

Search completed: November 5, 2003, 20:16:46
Job time : 10.2884 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 20:15:12 ; Search time 18.9085 Seconds
(without alignments)
1462.395 Million cell updates/sec

Title: US-09-677-374-4

Perfect score: 832

Sequence: 1 MRGCLQGSSLLIISVFLVGC.....IYTACPDQGRWQVISTEK 161

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 644079 seqs, 171749292 residues

Total number of hits satisfying chosen parameters: 644079

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	832	100.0	161	12	US-10-261-446-4
2	832	100.0	256	12	US-10-261-446-6
3	815	98.0	162	12	US-10-261-446-2
4	127.5	15.3	182	15	US-10-127-032-169
5	112	13.5	20	12	US-10-261-446-16
6	83	10.0	423	12	US/10/244
7	83	10.0	423	15	US/10/013
8	83	10.0	423	15	US/10/150
9	82.5	9.9	71	15	US-10-091-724-5
10	79.5	9.6	618	10	US-09-925-300-1381
11	79.5	9.6	890	15	US-10-156-761-14378
12	78.5	9.4	208	12	US-10-029-386-33055
13	78.5	9.4	651	9	US-09-861-597-1
14	78.5	9.4	651	12	US-10-414-760-1
15	78	9.4	438	12	US-10-244-821-88

16	77.5	9.3	274	15	US-10-156-761-9107	Sequence 9107, Ap
17	77.5	9.3	285	10	US-09-978-295A-28	Sequence 28, Appl
18	77.5	9.3	285	10	US-09-978-697-28	Sequence 28, Appl
19	77.5	9.3	285	10	US-09-978-192A-28	Sequence 28, Appl
20	77.5	9.3	285	10	US-09-999-832A-28	Sequence 28, Appl
21	77.5	9.3	285	11	US-09-978-189-28	Sequence 28, Appl
22	77.5	9.3	285	11	US-09-978-608A-28	Sequence 28, Appl
23	77.5	9.3	285	11	US-09-978-585A-28	Sequence 28, Appl
24	77.5	9.3	285	11	US-09-978-191A-28	Sequence 28, Appl
25	77.5	9.3	285	11	US-09-978-403A-28	Sequence 28, Appl
26	77.5	9.3	285	11	US-09-978-564A-28	Sequence 28, Appl
27	77.5	9.3	285	11	US-09-999-833A-28	Sequence 28, Appl
28	77.5	9.3	285	11	US-09-981-815A-28	Sequence 28, Appl
29	77.5	9.3	285	11	US-09-978-824-28	Sequence 28, Appl
30	77.5	9.3	285	11	US-09-918-585A-28	Sequence 28, Appl
31	77.5	9.3	285	11	US-09-978-423A-28	Sequence 28, Appl
32	77.5	9.3	285	11	US-09-978-193A-28	Sequence 28, Appl
33	77.5	9.3	285	11	US-09-999-830A-28	Sequence 28, Appl
34	77.5	9.3	285	11	US-09-978-757A-28	Sequence 28, Appl
35	77.5	9.3	285	11	US-09-978-187B-28	Sequence 28, Appl
36	77.5	9.3	285	11	US-09-978-643A-28	Sequence 28, Appl
37	77.5	9.3	285	12	US-09-978-375A-28	Sequence 28, Appl
38	77.5	9.3	285	12	US-09-978-188A-28	Sequence 28, Appl
39	77.5	9.3	285	12	US-09-978-298A-28	Sequence 28, Appl
40	77.5	9.3	285	12	US-10-143-031A-28	Sequence 28, Appl
41	77.5	9.3	285	12	US-10-002-967A-28	Sequence 28, Appl
42	77.5	9.3	285	12	US-10-017-083A-28	Sequence 28, Appl
43	77.5	9.3	285	12	US-10-143-030A-28	Sequence 28, Appl
44	77.5	9.3	285	12	US-10-199-672-4	Sequence 4, Appl
45	77.5	9.3	285	12	US-10-187-749-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1

US-10-261-446-4
; Sequence 4, Application US/10261446
; Publication No. US20030165526A1
; GENERAL INFORMATION:
; APPLICANT: Kuzyk, Michael A.
; APPLICANT: Burian, Jan
; APPLICANT: Kay, William W.
; APPLICANT: Thornton, Julian C.
; TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY IN FISH AGAINST
; TITLE OF INVENTION: RICKETTSIAL DISEASES, AND ASSOCIATED PREVENTATIVE THERAPY
; FILE REFERENCE: 4616-64094
; CURRENT APPLICATION NUMBER: US/10/261,446
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US 09/677,374
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/154,437
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Piscirickettsia salmonis
US-10-261-446-4

Query Match	100.0%	Score 832;	DB 12;	Length 161;
Best Local Similarity	100.0%	Pred. No. 1.1e-79;	Mismatches 0;	Indels 0; Gaps 0;
Matches 161;	Conservative 0;			
QY	1	MRGCLQGSSLLIISVFLVGC	1	MRGCLQGSSLLIISVFLVGC
Db	1	MRGCLQGSSLLIISVFLVGC	1	MRGCLQGSSLLIISVFLVGC
QY	61	LGGLIGSKIGSQMDQDKIKNLS	61	LGGLIGSKIGSQMDQDKIKNLS
Db	61	LGGLIGSKIGSQMDQDKIKNLS	61	LGGLIGSKIGSQMDQDKIKNLS

Qy 121 RRQYCRFFQKAMIAGQKEIYGTACPPDGRWQVISTEK 161
Dd 121 RRQYCRFFQKAMIAGQKEIYGTACPPDGRWQVISTEK 161

RESULT 2
US-10-261-446-6
; Sequence 6, Application US/10261446
; Publication No. US2003016526A1
; GENERAL INFORMATION:
; APPLICANT: Kuzyk, Michael A.
; APPLICANT: Burian, Jan
; APPLICANT: Kay, William W.
; APPLICANT: Thornton, Julian C.
; TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY IN FISH AGAINST
; TITLE OF INVENTION: RICKETTSIAL DISEASES, AND ASSOCIATED PREVENTATIVE THERAPY
; FILE REFERENCE: 4616-64094
; CURRENT APPLICATION NUMBER: US/10/261,446
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US 09/677,374
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/154,437
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Piscirickettsia salmonis
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (-95)..(-1)
US-10-261-446-6

RESULT 3
US-10-261-446-2
Sequence 2, Application US/10261446
Publication No. US20030165526A1
GENERAL INFORMATION:
APPLICANT: Kuzyk, Michael A.
APPLICANT: Burian, Jan
APPLICANT: Kay, William W.
APPLICANT: Thornton, Julian C.
TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY IN FISH AGAINST
TITLE OF INVENTION: RICKETTSIAL DISEASES, AND ASSOCIATED PREVENTATIVE THERAPY
FILE REFERENCE: 4616-64094
CURRENT APPLICATION NUMBER: US/10/261,446
CURRENT FILING DATE: 2002-09-30
PRIOR APPLICATION NUMBER: US 09/677,374
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/154,437
PRIOR FILING DATE: 1999-09-17
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 162

```

; TYPE: PRT
; ORGANISM: Piscirickettsia salmonis
US-10-261-446-2

Query Match          98.0%; Score 815; DB 12; Length 162;
Best Local Similarity 98.8%; Pred. No. 6.8e-78;
Matches 158; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY  2  RGCLOGSSLIISVFLVGCANFSRQEVGAATGAIVGGVAGQLFGKSGSRVSMaIGGAVL 61
      |||||
DB   3  RGCLOGSSLIISVFLVGCANFSRQEVGAATGAIVGGVAGQLFGKSGSRVSMaIGGAVL 62
      |||||

QY  62  GGLIGSKIGSMDDQDKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQER 121
      |||||
DB   63  GGLIGSKIGSMDDQDKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQER 122
      |||||

QY  122  RQCYCFEFOQKAMIGAKQOEIYGTACPPDGRWQVISTEK 161
      |||||
DB   123  RQCYCFEFOQKAMIGAKQOEIYGTACRQPDGRWQVISTEK 162
      |||||

RESULT 4
US-10-127-032-169
; Sequence 169, Application US/10127032
; Publication No. US20030113742A1
; GENERAL INFORMATION:
; APPLICANT: Whiteley, Marvin
; APPLICANT: Bangera, M. Gita
; APPLICANT: Lory, Stephen
; APPLICANT: Greenberg, Everett Peter
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF
; TITLE OF INVENTION: BIOFILM FORMATION
; FILE REFERENCE: UIZ-070CP
; CURRENT APPLICATION NUMBER: US/10/127,032
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/285,190
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: US 60/344,142
; PRIOR FILING DATE: 2001-10-24
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 169
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-127-032-169

Query Match          15.3%; Score 127.5; DB 15; Length 182;
Best Local Similarity 34.8%; Pred. No. 1.8e-05;
Matches 33; Conservative 13; Mismatches 28; Indels 19; Gaps 2;

QY  27  QEVGAATGAVVGGVAGQLFGKSGSRVSMaIGGAVLGLIGSKIGSQSMDDQDKI----- 79
DB   70  QIAGTAGVAVVGLLNGIGGGTGKXIATVAGAVGGYAGNKVQOEGNQERDVTYTTTETRC 129
      |||||

QY  80  -KLNQSLKLV-----KAGQVTRWRNP 99
      ::|||
DB   130  STVHDSSEKVVGYDVKYMLDGKAGQIRMERDP 161
      ::|||

RESULT 5
US-10-261-446-16
; Sequence 16, Application US/10261446
; Publication No. US20030165526A1
; GENERAL INFORMATION:
; APPLICANT: Kuzyk, Michael A.
; APPLICANT: Burian, Jan
; APPLICANT: Kay, William W.
; APPLICANT: Thornton, Julian C.
; TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY IN FISH AGAINST
; TITLE OF INVENTION: RICKETTSIAL DISEASES, AND ASSOCIATED PREVENTATIVE THERA
; FILE REFERENCE: 4616-64094
; CURRENT APPLICATION NUMBER: US/10/261.446

```

; CURRENT FILING DATE: 2002-09-30
 ; PRIOR APPLICATION NUMBER: US 09/677,374
 ; PRIOR FILING DATE: 2000-09-15
 ; PRIOR APPLICATION NUMBER: US 60/154,437
 ; PRIOR FILING DATE: 1999-09-17
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 16
 ; LENGTH: 20
 ; TYPE: PRT
 ; ORGANISM: *Piscirickettsia salmonis*
 US-10-261-446-16

Query Match 13.5%; Score 112; DB 12; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.2e-05;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 PVRTYQRYNKQRRQYCRE 128
 |||||
 DB 1 PVRTYQRYNKQRRQYCRE 20

RESULT 6
 US/10/244
 ; Sequence 8, Application US/10244821
 ; Publication No. US20030143233A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Goshorn, Stephen Charles
 ; APPLICANT: Graves, Scott Stoll
 ; APPLICANT: Schultz, Joanne Elaine
 ; APPLICANT: Lin, Yukang
 ; APPLICANT: Sanderson, James Allen
 ; APPLICANT: Reno, John M.
 ; APPLICANT: Dearstynne, Erica A.
 ; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
 ; FILE REFERENCE: 690022.547C3
 ; CURRENT FILING DATE: 2002-09-16
 ; NUMBER OF SEQ ID NOS: 92
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 8
 ; LENGTH: 423
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Predicted amino acid sequence of B9E9 single chain antibody-
 US/10/244,821-8

Query Match 10.0%; Score 83; DB 12; Length 423;
 Best Local Similarity 21.7%; Pred. No. 2.6;
 Matches 33; Conservative 24; Mismatches 67; Indels 28; Gaps 5;

QY 30 GAATGAVVGGVAGQLFGKSGRVSMAIGGAVLGGIGSKIGSQMDQDKIKLQN----- 83
 |||||
 DB 134 GSGGGSGGGSG---GGSSDIVLSQSPAILSPGKVTMTCRASSSVSYMHVYQKP 190
 QY 84 -----SLEKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQRRQYCREP--QOK 132
 |||||
 DB 191 GSSPKPWIYATNSLGSVPFSGSGTSYSL-----TISRVEADAATYYCQWISNPP 246
 QY 133 AMIAGQKQEI-----YGTACPDGRWQVISTE 160
 |||||
 DB 247 TFGAGTKLELKSXSGSADPSKDSKAQVSAE 278

RESULT 7
 US/10/013
 ; Sequence 8, Application US/10013173
 ; Publication No. US20030095977A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Goshorn, Stephen C.
 ; APPLICANT: Graves, Scott Stoll

; APPLICANT: Schultz, Joanne Elaine
 ; APPLICANT: Lin, Yukang
 ; APPLICANT: Sanderson, James A.
 ; APPLICANT: Reno, John M.
 ; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
 ; FILE REFERENCE: 690022.547C1
 ; CURRENT FILING DATE: 2001-12-07
 ; NUMBER OF SEQ ID NOS: 69
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 8
 ; LENGTH: 423
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Predicted amino acid sequence of B9E9 single chain antibody-
 US/10/013,173-8

Query Match 10.0%; Score 83; DB 15; Length 423;
 Best Local Similarity 21.7%; Pred. No. 2.6;
 Matches 33; Conservative 24; Mismatches 67; Indels 28; Gaps 5;
 QY 30 GAATGAVVGGVAGQLFGKSGRVSMAIGGAVLGGIGSKIGSQMDQDKIKLQN----- 83
 |||||
 DB 134 GSGGGSGGGSG---GGSSDIVLSQSPAILSPGKVTMTCRASSSVSYMHVYQKP 190
 QY 84 -----SLEKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQRRQYCREP--QOK 132
 |||||
 DB 191 GSSPKPWIYATNSLGSVPFSGSGTSYSL-----TISRVEADAATYYCQWISNPP 246
 QY 133 AMIAGQKQEI-----YGTACPDGRWQVISTE 160
 |||||
 DB 247 TFGAGTKLELKSXSGSADPSKDSKAQVSAE 278

RESULT 8
 US/10/150
 ; Sequence 8, Application US/10150762
 ; Publication No. US20030103948A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Goshorn, Stephen C.
 ; APPLICANT: Graves, Scott S.
 ; APPLICANT: Schultz, Joanne E.
 ; APPLICANT: Lin, Yukang
 ; APPLICANT: Sanderson, James A.
 ; APPLICANT: Reno, John M.
 ; APPLICANT: Dearstynne, Erica A.
 ; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
 ; FILE REFERENCE: 690022.547C2
 ; CURRENT FILING DATE: 2002-05-17
 ; NUMBER OF SEQ ID NOS: 90
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 8
 ; LENGTH: 423
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Predicted amino acid sequence of B9E9 single chain antibody-
 US/10/150,762-8

Query Match 10.0%; Score 83; DB 15; Length 423;
 Best Local Similarity 21.7%; Pred. No. 2.6;
 Matches 33; Conservative 24; Mismatches 67; Indels 28; Gaps 5;
 QY 30 GAATGAVVGGVAGQLFGKSGRVSMAIGGAVLGGIGSKIGSQMDQDKIKLQN----- 83
 |||||
 DB 134 GSGGGSGGGSG---GGSSDIVLSQSPAILSPGKVTMTCRASSSVSYMHVYQKP 190
 QY 84 -----SLEKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQRRQYCREP--QOK 132
 |||||

Db 191 GSSPKWIVATNSLGSVPARTSGSGTSYSL-----TISRVEAEDAATYYQQWISNPP 246

QY 133 AMAGKQKEI-----YGTACPDQPDGRWOVISTE 160

Db 247 TFCAGTKLELKSGSGSADPSKSKAQVGAEE 278

RESULT 9

US-10-091-724-5

Sequence 5, Application US/10091724

Publication No. US20030105310A1

GENERAL INFORMATION:

APPLICANT: Children's Medical Center Corporation

APPLICANT: Ashkar, Samy

TITLE OF INVENTION: Method to Screen Peptide Libraries Using Minicell Display

FILE REFERENCE: CMCC 820

CURRENT APPLICATION NUMBER: US/10/091,724

CURRENT FILING DATE: 2002-03-06

PRIOR APPLICATION NUMBER: US 60/306,946

PRIOR FILING DATE: 2001-07-20

PRIOR APPLICATION NUMBER: 60/274,039

PRIOR FILING DATE: 2001-03-07

NUMBER OF SEQ ID NOS: 53

SOFTWARE: PatentIn version 3.1

SEQ ID NO 5

LENGTH: 71

TYPE: PRT

ORGANISM: Artificial sequence

FEATURE:

OTHER INFORMATION: First 71 amino acids of the 17 K antigen of Rickettsia rickettsii

US-10-091-724-5

Query Match 9.9%; Score 82.5; DB 15; Length 71;

Best Local Similarity 32.4%; Pred. No. 0.28; Indels 24; Gaps 4;

Matches 22; Conservative 15; Mismatches 24; Indels 7; Gaps 4;

QY 8 SSLIIISV---FLVGC--AQNFSRQEVGAATGAVGVAGQLFGKSGSRVSMAG-GAVL 61

Db 5 SKINIALATSMALACNCPGGMKQGTGTLGGAGALLGQFGKSGQL-VGVGVGALL 63

QY 62 GGLIGSKI 69

Db 64 GAVLGGQI 71

RESULT 10

US-09-925-300-1381

Sequence 1381, Application US/09925300

Patent No. US20020151681A1

GENERAL INFORMATION:

APPLICANT: Craig Rosen,

APPLICANT: Steve Ruben

TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies

FILE REFERENCE: PA101

CURRENT APPLICATION NUMBER: US/09/925,300

CURRENT FILING DATE: 2001-08-10

PRIOR APPLICATION NUMBER: PCT/US00/05988

PRIOR FILING DATE: 2000-03-08

PRIOR APPLICATION NUMBER: 60/124,270

PRIOR FILING DATE: 1999-03-12

NUMBER OF SEQ ID NOS: 1890

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 1381

LENGTH: 618

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SITE

LOCATION: (507)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE

LOCATION: (524)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

NAME/KEY: SITE

LOCATION: (562)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-925-300-1381

Query Match 9.6%; Score 79.5; DB 10; Length 618;

Best Local Similarity 28.7%; Pred. No. 9.8;

Matches 41; Conservative 15; Mismatches 54; Indels 33; Gaps 10;

QY 3 GCLQGSSLIISVFP--LVGCAQNFSRQEVGAAT--GAVVGVAGQLFGKSGS-GRVSM--- 54

Db 76 GCFGGSS-----GGYGGLGGFGGSGFRGSGSGSGSGSGSGSGSGSGSGSGSGSG 131

QY 55 AIGGAVLGLGSGIKGQS---MDQDKI-----KLNQSLKRYKA-----GVTR 95

Db 132 GFGGGFGGFGGFGGFGGGLSGNEKVTMQLNDRLASLYLDKVRALLESNVELEGIKE 191

QY 96 WRNPDTGNSYSVEPVRTYQRYNK 118

Db 192 WVEKH-GNSHQEP-RDYSKYK 212

RESULT 11

US-10-156-761-14378

Sequence 14378, Application US/10156761

Publication No. US20030119018A1

GENERAL INFORMATION:

APPLICANT: OMURA, SATOSHI

APPLICANT: IKEDA, HARUO

APPLICANT: ISHIKAWA, JUN

APPLICANT: HORIKAWA, HIROSHI

APPLICANT: SHIBA, TADAYOSHI

APPLICANT: SAKAKI, YOSHIYUKI

APPLICANT: HATTORI, MASAHIRA

TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES

FILE REFERENCE: 249-262

CURRENT APPLICATION NUMBER: US/10/156,761

CURRENT FILING DATE: 2002-05-29

PRIOR APPLICATION NUMBER: JP 2001-204089

PRIOR FILING DATE: 2001-05-30

PRIOR APPLICATION NUMBER: JP 2001-272697

PRIOR FILING DATE: 2001-08-02

NUMBER OF SEQ ID NOS: 15109

SEQ ID NO 14378

LENGTH: 890

TYPE: PRT

ORGANISM: Streptomyces avermitilis

US-10-156-761-14378

Query Match 9.6%; Score 79.5; DB 15; Length 890;

Best Local Similarity 27.9%; Pred. No. 16;

Matches 24; Conservative 18; Mismatches 33; Indels 11; Gaps 2;

QY 1 MRGCLQGSSLIISVFP-----LVGCAQNFSRQEVGAATGAVGVAGQLFGKSGSRVS 53

Db 803 VRGRIQGGRPAPVALFTTVGKPLTVIATNAAERGLKAGLDLVRAAAKTLGGGGGGRPD 862

QY 54 MAIGGAVLGLGSGIKGSGMDQDKI 79

Db 863 VAQG-----GGQNPAAGDAIDAVERL 884

RESULT 12

US-10-029-386-33055

Sequence 33055, Application US/10029386

Publication No. US20030194704A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G

TITLE OF INVENTION: EXPRESSION ANALYSIS TWO

FILE REFERENCE: AEOMICA-X-2

```

; GENERAL INFORMATION:
; APPLICANT: Yang, Jianjun G.
; TITLE OF INVENTION: Production of Silk-Like Proteins in Plants
; FILE REFERENCE: BC1014 US NA
; CURRENT APPLICATION NUMBER: US/10/414,760
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: US/09/863,859
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: 60/206968
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Nephila clavipes
; US-10-414-760-1

Query Match          9.4%; Score 78.5; DB 12; Length 651;
Best Local Similarity 35.3%; Pred. No. 13;
Matches 24; Conservative 4; Mismatches 17; Indels 23; Gaps 2
QY      27  QEYGAATGAVVG---GVAGQ-----LFGKSGSRVSNIAIGGAVLGG 63
Db      491  QGAGAAAAAAGVAGQEGIRGQAGAGCGYGGIGSGSGRGGLGGQAGAAAAAGAGGQGG 550
QY      64  LIGSKITGQ 71
Db      551  LGGQGHGQ 558

```

```

RESULT 15
US-10-244-821-88
; Sequence 88, Application US/10244821
; Publication No. US20030143233A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen Charles
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James Allen
; APPLICANT: Reno, John M.
; APPLICANT: Dearstyne, Erica A.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 690022.547C3
; CURRENT APPLICATION NUMBER: US/10/244,821
; CURRENT FILING DATE: 2002-09-16
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 88
; LENGTH: 438
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-244-821-88

```

Query Match	9.4%	Score 78;	DB 12;	Length 438;
Best Local Similarity	21.1%;	Pred. No. 9;		
Matches	32;	Conservative 24;	Mismatches 68;	Indels 28; Gaps 5

Qy	30	GAATGAVVGGVAGOLFEGKSGGRVSMAGGGAVLGLGLGSKIGSQMDQDKIKLQ-----	83
Db	149	GGGGGGGGGGSG-----GGGSSQIVLTQSPAIMSASPGKEKVTITCSASSISYMEWFOQKP	205
Qy	84	-----SLEKVKAGQVTRWRNPDTGNSYVSVEPVRYORYNKQRQQYCRFFQQKAM	134
Db	206	GTSPKLWIYTTNSLQSGVPARFSGSGSGTSYSL-----TISRMEADATYYCHQRSTYPL	261
Qy	135	I--AGQKEI-----YGTACQPQDGRWQVISTE	160
Db	262	TFSGTGKLELKSQSGSGSADPSKDSKAQVSAE	293

QY 135 I--AGUQEJ-----IGIACFPDGRWVVISLE 160
 :|:|:| :|:|:| :|:|:|
Db 262 TFGSGTKLELKSSGSGSADFSKDSKAQVSAAE 293

Search completed: November 5, 2003, 20:27:19
Job time : 18.9085 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 20:10:11 ; Search time 11.4007 Seconds
(without alignments)
1358.069 Million cell updates/sec

Title: US-09-677-374-4

Perfect score: 832

Sequence: 1 MRGCLQSSLLIIISVFLVGC.....IYTACPOPGRGWVISTEK 161

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	288	34.6	159	D33971	rickettsial common
2	285	34.3	159	B33971	Rickettsial common
3	285	34.3	159	D33971	Rickettsial common
4	285	34.3	159	G97860	17K surface antigen
5	276.5	33.2	159	D33971	Rickettsial common
6	276	33.2	159	A25972	17K antigen precursor
7	127.5	15.3	182	D83169	conserved hypothet
8	120.5	14.5	131	A13418	17K surface antigen
9	112.5	13.5	155	S23787	outer membrane lip
10	111.5	13.4	154	B83514	conserved hypothet
11	109	13.1	179	AE0644	probable secreted
12	109	13.1	232	G87629	hypothetical prote
13	109	13.1	257	B82837	conserved hypothet
14	107.5	12.9	142	A26696	lipa protein limpo
15	107.5	12.9	155	AF0289	probable lipoprote
16	107	12.9	125	D97478	lipa protein limpo
17	105.5	12.7	155	A10693	outer membrane lip
18	104.5	12.6	155	C64921	outer membrane lip
19	104.5	12.6	155	F90922	probable outer mem
20	104.5	12.6	155	B85771	probable outer mem
21	104	12.5	139	S58234	lipa protein - Rhi
22	103	12.4	155	AG0443	outer membrane lip
23	103	12.4	232	AD3350	outer membrane pro
24	101	12.1	155	I64130	PAL cross-reacting
25	100.5	12.1	257	F87413	hypothetical prote
26	100	12.0	179	D85674	hypothetical prote
27	100	12.0	179	H90814	hypothetical prote
28	100	12.0	179	C64855	ycfJ protein - Esc
29	99.5	12.0	304	H83636	hypothetical prote

RESULT 1
D33971
rickettsial common antigen precursor - Rickettsia prowazekii
N:Alternate names: 17kD surface antigen; outer membrane protein (omp); RP833
C:Species: Rickettsia prowazekii
C:Date: 16-Mar-1990 #sequence_revision 16-Mar-1990 #text_change 03-Nov-2000
C:Accession: D33971; B71645
R:Anderson, B.E.; Tzianabos, T.
J. Bacteriol. 171, 5199-5201, 1989
A:Title: Comparative sequence analysis of a genus-common rickettsial antigen gene.
A:Reference number: A33971; MUID:89359171; PMID:2768201
A:Accession: D33971
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-159 <AND>
A:Cross-references: GB:M28482; NID:g152461
A>Note: the sequence in GenBank entry RIRANTI7KC, release 109.0, (PID:g152462) omits the
R:Anderson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichert-Ponten, T.; Alsmark, U.
Nature 396, 133-140, 1998
A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A:Reference number: A71630; MUID:99039499; PMID:9823893
A:Accession: B71645
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-159 <AN2>
A:Cross-references: GB:AJ235273; GB:AJ235269; NID:g3861237; PIDN:CAA15258.1; PID:g3861335
A:Experimental source: strain Madrid E
C:Genetics:
A:Gene: omp; RP833
C:Superfamily: rickettsial common antigen
C:Keywords: surface antigen

ALIGNMENTS

Query Match 34.6%; Score 288; DB 2; Length 159;
Best Local Similarity 37.9%; Pred. No. 1.6e-18;
Matches 61; Conservative 32; Mismatches 52; Indels 16; Gaps 6;
QY 8 SLLIIISV---FLVGC--AQNFSRQEVGAATGAVGVAGQLFGKSGRVSMAIG-GAVL 61
DB 5 SKIMTIALAASMLQACNCGSGMKNQGTGLLGGAGGALLGSGQGGKGL-VGVGVGALL 63
QY 62 GGLIGSKTIGSQNDQDK---IKLNQSLKVKAGQVTRWRNPDTGNSYSVEPRTYQRXN 117
DB 64 GAVLGGQIGASNDQDRLELTSORALESPSGSNIEWRNPDNHGVITPNKY---- 119
QY 118 KQERQQYCRFFQQXAMTAGQKEIYGFACPDGRGWVIS 158
DB 120 -RNSAGQYCREYTVIIGKQKQTYGNACRQPDGQWQVN 159

RESULT 2

B33971
Rickettsial common antigen precursor - Rickettsia conorii

C:Species: Rickettsia conorii
 C>Date: 16-Mar-1990 #sequence_revision 16-Mar-1990 #text_change 20-Aug-1999
 C:Accession: B33971
 R:Anderson, B.E.; Tzianabos, T.
 J. Bacteriol. 171, 5199-5201, 1989
 A:Title: Comparative sequence analysis of a genus-common rickettsial antigen gene.
 A:Reference number: A33971; MUID:89359171; PMID:2768201
 A:Accession: B33971
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-159 <AND>
 A:Cross-references: GB:M28479; NID:gl52463; PIDN:AAA26379.1; PID:gl52464
 C:Superfamily: rickettsial common antigen

Query Match 34.3%; Score 285; DB 2; Length 159;
 Best Local Similarity 37.9%; Pred. No. 3e-18;
 Matches 61; Conservative 31; Mismatches 53; Indels 16; Gaps 6;
 QY 8 SSLIIISV---FLVGC--AQNFSRQEVGAATGAVGVGAGQLFGKSGRVSMAIG-GAVL 61
 DB 5 SKMIIALATSMLOACNPGGNKQGTGTLGGAGALLGSQFGKGGQL-VGVGVGALL 63
 QY 62 GGLIGSKIQSMDDQDK----IKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYN 117
 DB 64 GAVLGGQIGAGMDEQDRRLAELTSQRALETAPSGSNVWRNPDNNGYVTPNKTY---- 119
 QY 118 KQRRQYCRFFQOKAMIAQKQEIYGTACPDGRWQVIS 158
 DB 120 -RNSTGYCREYQTQTVVIGKQKQAYGNACRQPDGQWQVNV 159

RESULT 3
 A33971
 Rickettsial common antigen precursor - Rickettsia rickettsii
 C:Species: Rickettsia rickettsii
 C>Date: 16-Mar-1990 #sequence_revision 16-Mar-1990 #text_change 20-Aug-1999
 C:Accession: A33971
 R:Anderson, B.E.; Tzianabos, T.
 J. Bacteriol. 171, 5199-5201, 1989
 A:Title: Comparative sequence analysis of a genus-common rickettsial antigen gene.
 A:Reference number: A33971; MUID:89359171; PMID:2768201
 A:Accession: A33971
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-159 <AND>
 A:Cross-references: GB:M28480; NID:gl52457; PIDN:AAA26376.1; PID:gl52458
 C:Superfamily: rickettsial common antigen

Query Match 34.3%; Score 285; DB 2; Length 159;
 Best Local Similarity 37.9%; Pred. No. 3e-18;
 Matches 61; Conservative 31; Mismatches 53; Indels 16; Gaps 6;
 QY 8 SSLIIISV---FLVGC--AQNFSRQEVGAATGAVGVGAGQLFGKSGRVSMAIG-GAVL 61
 DB 5 SKMIIALATSMLOACNPGGNKQGTGTLGGAGALLGSQFGKGGQL-VGVGVGALL 63
 QY 62 GGLIGSKIQSMDDQDK----IKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYN 117
 DB 64 GAVLGGQIGAGMDEQDRRLAELTSQRALETAPSGSNVWRNPDNNGYVTPNKTY---- 119
 QY 118 KQRRQYCRFFQOKAMIAQKQEIYGTACPDGRWQVIS 158
 DB 120 -RNSTGYCREYQTQTVVIGKQKQAYGNACRQPDGQWQVNV 159

RESULT 4
 G97860
 17K surface antigen precursor [imported] - Rickettsia conorii (strain Malish 7)
 C:Species: Rickettsia conorii
 C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 22-Oct-2001
 C:Accession: G97860
 R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; R
 Science 293, 2093-2098, 2001

A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
 A:Reference number: A97700; MUID:21442074; PMID:11557893
 A:Accession: G97860
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-159 <KOR>
 A:Cross-references: GB:AB006914; PIDN:AAO3825.1; PID:gl5260425; GSPDB:GN00173
 C:Genetics:
 A:Gene: omp
 C:Superfamily: rickettsial common antigen

Query Match 34.3%; Score 285; DB 2; Length 159;
 Best Local Similarity 37.9%; Pred. No. 3e-18;
 Matches 61; Conservative 31; Mismatches 53; Indels 16; Gaps 6;
 QY 8 SSLIIISV---FLVGC--AQNFSRQEVGAATGAVGVGAGQLFGKSGRVSMAIG-GAVL 61
 DB 5 SKMIIALATSMLOACNPGGNKQGTGTLGGAGALLGSQFGKGGQL-VGVGVGALL 63
 QY 62 GGLIGSKIQSMDDQDK----IKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYN 117
 DB 64 GAVLGGQIGAGMDEQDRRLAELTSQRALETAPSGSNVWRNPDNNGYVTPNKTY---- 119
 QY 118 KQRRQYCRFFQOKAMIAQKQEIYGTACPDGRWQVIS 158
 DB 120 -RNSTGYCREYQTQTVVIGKQKQAYGNACRQPDGQWQVNV 159

RESULT 5
 C33971
 Rickettsial common antigen precursor - Rickettsia typhi
 C:Species: Rickettsia typhi
 C>Date: 16-Mar-1990 #sequence_revision 16-Mar-1990 #text_change 20-Aug-1999
 C:Accession: C33971
 R:Anderson, B.E.; Tzianabos, T.
 J. Bacteriol. 171, 5199-5201, 1989
 A:Title: Comparative sequence analysis of a genus-common rickettsial antigen gene.
 A:Reference number: A33971; MUID:89359171; PMID:2768201
 A:Accession: C33971
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-159 <AND>
 A:Cross-references: GB:M28481; NID:gl52459; PIDN:AAA26377.1; PID:gl52460
 C:Superfamily: rickettsial common antigen

Query Match 33.2%; Score 276.5; DB 2; Length 159;
 Best Local Similarity 39.6%; Pred. No. 1.7e-17;
 Matches 55; Conservative 27; Mismatches 46; Indels 11; Gaps 4;
 QY 25 SRQEVGAATGAVGVGAGQLFGKSGRVSMAIG-GAVLGLIGSKIQSMDDQDK---I 79
 DB 27 NKQGTGTLGGAGALLGSQFGKGGQL-VGVGVGALLGVGGQIGASLDBQDRKLEL 85
 QY 80 KLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKOERRQYCRFFQOKAMIAQK 139
 DB 86 TSQRALESAPSGSNTEWRNPDNNGYVTPNKTY-----RNSTGYCREYQTQTVVIGGKQ 140
 QY 140 QEIYGTACPDGRWQVIS 158
 DB 141 QTTYGNACRQPDGQWQVNV 159

RESULT 6
 A25972
 17K antigen precursor - Rickettsia rickettsii
 C:Species: Rickettsia rickettsii
 C>Date: 21-May-1988 #sequence_revision 21-May-1988 #text_change 20-Aug-1999
 C:Accession: A25972
 R:Anderson, B.E.; Regnery, R.L.; Carlone, G.M.; Tzianabos, T.; McDade, J.E.; Fu, Z.Y.;
 J. Bacteriol. 169, 2385-2390, 1987
 A:Title: Sequence analysis of the 17-kilodalton-antigen gene from Rickettsia rickettsii
 A:Reference number: A25972; MUID:87222152; PMID:3108232
 A:Accession: A25972

A:Molecule type: DNA
A:Residues: 1-159 <AND>
A:Cross-references: GB:M16486; NID:g152467; PIDN:AAA26381.1; PID:g152468
C:Superfamily: rickettsial common antigen

Query Match 33.2%; Score 276; DB 2; Length 159;
Best Local Similarity 37.3%; Pred. No. 1.9e-17;
Matches 60; Conservative 31; Mismatches 54; Indels 16; Gaps 6;

QY 8 SLLIISV---FLVGC--AQNFSDREVGAATGAVGVAGQLFGKSGRVSNALIG-GAVL 61
DB 5 SKIMIALATSMLOACNGPFGNKKGTGLLGGAGALLGQFGKQQL-VGVGVGALL 63

QY 62 GGLIGSKTQSQMDQDK----IKNQSLKVKAGQVTRWRNPDTCNSYSVBPVRYRYN 117
DB 64 GAVLGGQIGAGNDEDRRLAELTSQALETAPSGSNVWRPNDCNGYGVTPNKTY---- 119

QY 118 KQRRQQYCRFPQQKAMTAGQKQEIYGTACPODGRWQVIS 158
DB 120 -RNSTGQYCREYTQTWVIGGKQKAYGACRQPDQWQVNV 159

RESULT 7
DB3169
conserved hypothetical protein PA3819 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 01-Mar-2002
C:Accession: DB3169
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bradman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: DB3169
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-182 <STO>
A:Cross-references: GB:AE004799; GB:AE004091; NID:g9949981; PIDN:AAG07206.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
C:Superfamily: PAL cross-reacting lipoprotein

Query Match 15.3%; Score 127.5; DB 2; Length 182;
Best Local Similarity 34.8%; Pred. No. 0.00037;
Matches 32; Conservative 13; Mismatches 28; Indels 19; Gaps 2;

QY 27 QEVGAATGAVGVAGQLFGKSGRVSNALIGGAVLGLIGSKIGOSMDQDKI----- 79
DB 70 QIAGTAIGAVVGLLGNQIGGTGKKIATVAGAVGGYAGNKVQSGMQERDTYTTETRC 129

QY 80 -KLNQSLRKV-----KAGQVTRWRNP 99
DB 130 STVHDSSEKVGVDYVKYMLDGKAGQIRMERDP 161

RESULT 8
AI3418
17K surface antigen precursor [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: AI3418
R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, E.; Mazur, M.; Goltzman, E.; Selkov, E.; Elizer, P.H.; Hagius, S.; O'Callaghan, D.; Letessier, Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AJ3252; PMID:11756688
A:Accession: AI3418
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-131 <KUR>
A:Cross-references: GB:AE008917; PIDN:AAL52516.1; PID:g17983328; GSPDB:GN00190

A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI1335
A:Map position: 1

Query Match 14.5%; Score 120.5; DB 2; Length 131;
Best Local Similarity 31.7%; Pred. No. 0.0011;
Matches 39; Conservative 17; Mismatches 44; Indels 23; Gaps 7;

QY 46 GKSGSRVSNALIG-----AVLGG-GLIGSKIGQ--SMDQDKIKLN-QSLEKVKAG 91
DB 14 GKSGGFPF--LGGSSQRPETWLLASLGNLFGNSASQLSAADRRKALAEYRALYSFAG 71

QY 92 QVTRWRNPDTCNSYSVBPVRYRYNQRRQQYCRFPQQKAMTAGQKQEIYGTACPOPD 151
DB 72 KSVLWSGAGS-NAGDVTAAPYQ-----VGSQNCROYSHSFTTIGDQDQTVRGTAACRNP 124

QY 152 GRW 154
DB 125 GSW 127

RESULT 9
S23787
outer membrane lipoprotein precursor - Yersinia enterocolitica
C:Species: Yersinia enterocolitica
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 29-Sep-1999
C:Accession: S23787
R:Baeumler, A.J.; Hantke, K.
J. Bacteriol. 174, 1029-1035, 1992
A:Title: A lipoprotein of Yersinia enterocolitica facilitates ferrioxamine uptake in Escherichia coli
A:Reference number: S23786; MUID:92121089; PMID:1732192
A:Accession: S23787
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-155 <BAB>
A:Cross-references: NID:X60448; NID:g48577; PIDN:CAA42977.1; PID:g48579
C:Superfamily: PAL cross-reacting lipoprotein

Query Match 13.5%; Score 112.5; DB 2; Length 155;
Best Local Similarity 24.7%; Pred. No. 0.0067;
Matches 37; Conservative 19; Mismatches 43; Indels 51; Gaps 4;

QY 9 SLLIISVFLVCAON-----FSRQE-----VGAAT 33
DB 7 AVATAAVTLTGCAANNLTSLGDFSAQKQVQTVTYTGLTSVRPVTIOGGDDNNVGAIG 66

QY 34 GAVGVGAVAGQLFGKSGRVSNALIGGAVLGLIGSKIGOSMDQDKIKL----- 81
DB 67 GAVLGGPLGNTVGGTGRSLATAAGAVAGGAGQVQGMNRTDGVQLEVRKDDGTTLV 126

QY 82 --NQSLKVKAGQVTRWRNPDTCNSYSVEP 109
DB 127 VQKQPTRFVQ--RVMLASSGSTVTVSP 154

RESULT 10
B83514
conserved hypothetical protein PA1053 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: B83514
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bradman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: B83514
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-154 <STO>
A:Cross-references: GB:AE004537; GB:AE004091; NID:g9946960; PIDN:AAG04442.1; GSPDB:GN001

A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PAI053
 C:Superfamily: PAL cross-reacting lipoprotein

Query Match 13.4%; Score 111.5; DB 2; Length 154;
 Best Local Similarity 25.7%; Pred. No. 0.0082;
 Matches 35; Conservative 21; Mismatches 35; Indels 45; Gaps 4;

QY 5 LQSSLIITIS-----VFLVGC-----AQNFSROE----- 28
 DB 1 MKKSALIVASFAMALALGCGSSLTGDTYSREARTVQTVMGTIQAALRPVKIBGTKTP 60

QY 29 VGNATCAVGVAGQQLFGKSGRVSMAGVAVGLGLIGSKIQSMDQDKIKL----- 81
 DB 61 IGSIIAGAGVGVAGSANGGKGSYVAIIIGAVAGGLGAATBEGLTRTQGVSEITVREDDG 120

QY 82 --NQSLKVKAGQVTR 95
 DB 121 STRAVVQVDQGIIFR 136

RESULT 11
 AE0644
 probable secreted protein STY1252 [imported] - Salmonella enterica subsp. enterica serov
 C:Species: Salmonella enterica subsp. enterica serovar Typhi
 A:Note: this species has also been called Salmonella typhi
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C:Accession: AE0644
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
 th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
 S.; Moulé, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
 A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
 A:Reference number: AB0502; MUID:21534947; PMID:11677608
 A:Accession: AE0644
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-179 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD08336.1; PID:g16502381; GSPDB:GN00176
 C:Genetics:
 A:Gene: STY1252
 C:Superfamily: PAL cross-reacting lipoprotein

Query Match 13.1%; Score 109; DB 2; Length 179;
 Best Local Similarity 27.4%; Pred. No. 0.016;
 Matches 37; Conservative 16; Mismatches 50; Indels 32; Gaps 4;

QY 30 GAATGAVGVAGQQLFGKSGRVSMAGVAVGLGLIGSKIQSMDQDKIKLQSLKVK 89
 DB 73 GSVLGAAGVGVIGHQFGGGRGKDVATVVALGGYAGNQLQSGMQESD----- 120

QY 90 AGQVTRWRNPDTGNSYSVBPRTYRYNKRQYCRFQ---QKAMTAGQKQIYGP 146
 DB 121 -----TYTTTQRCCKTVYDKSERKMLGYDVTVYKIGDQDQKIRMDKP--GTQ 164

QY 147 CPQPDGRWQVISTEK 161
 DB 165 IPL-DGNGQLVLNKK 178

RESULT 12
 GB7629
 hypothetical protein CC3073 [imported] - Caulobacter crescentus
 C:Species: Caulobacter crescentus
 C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
 C:Accession: GB7629
 R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Laub, M.T.; DeBoy R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647
 A:Accession: GB7629
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-232 <STO>
 A:Cross-references: GB:AB005673; NID:g13424723; PIDN:AAK25035.1; GSPDB:GN00148
 C:Genetics:
 A:Gene: CC3073

Query Match 13.1%; Score 109; DB 2; Length 232;
 Best Local Similarity 33.7%; Pred. No. 0.021;
 Matches 32; Conservative 14; Mismatches 33; Indels 16; Gaps 2;

QY 17 LVGCAQNFSSROEVAATGAVGVAGQQLFGK---GSGRVSMAGVAVGLGLIGSKIQSM 73
 DB 81 VVGKASGKGEVAVGVAGLGAAGSNLAKNDQGTGTGAIGVAVGAGAGSLIGCKMQKSD 140

QY 74 DQDKIKLNQS-----LEKVKAGQVTR 95
 DB 141 AAQEVGGIYKSGFRYAQTVOAALPKVIEKQWTR 175

RESULT 13
 B82837
 conserved hypothetical protein XF0178 [imported] - Xylella fastidiosa (strain 9a5c)
 C:Species: Xylella fastidiosa
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
 C:Accession: B82837
 R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
 Nature 406, 151-157, 2000
 A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A:Reference number: A82515; MUID:20365717; PMID:10910347
 A:Note: for a complete list of authors see reference number A59328 below
 A:Accession: B82837
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-257 <SIN>
 A:Cross-references: GB:AE003872; GB:AE003849; NID:g9104975; PIDN:AAF82991.1; GSPDB:GN001
 A:Experimental source: strain 9a5c
 R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
 Briones, M.R.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
 as-Neto, E.; Docena, C.; El-Dorri, H.; Facinani, A.P.; Ferreira, A.J.S.
 submitted to GenBank, June 2000
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Kiege, J.E.; Kuramae, E.E.; Laig
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
 A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
 A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
 M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
 A:Reference number: A59328
 A:Contents: annotation
 C:Genetics:
 A:Gene: XF0178

Query Match 13.1%; Score 109; DB 2; Length 257;
 Best Local Similarity 47.5%; Pred. No. 0.024; Mismatches 12; Indels 0; Gaps 0;

QY 30 GAATGAVGVAGQQLFGKSGRVSMAGVAVGLGLIGSKI 69
 DB 105 GTAIGALIGLVGNQFGHNGRKAALTAAGAVAGGFIGNEV 144

RESULT 14
 AD2696
 lipA protein [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
 C:Species: Agrobacterium tumefaciens
 C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
 C:Accession: AD2696
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell

Search completed: November 5, 2003, 20:15:56
Job time : 11.4007 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 20:06:36 ; Search time 6.39551 Seconds
(without alignments)
1183.846 Million cell updates/sec

Title: US-09-677-374-4

Perfect score: 832

Sequence: 1 MRGCLQGSSLIISVFLVGC.....IYGTACPODGRWQVISTEK 161

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	288	34.6	159	1	17KD_RICPR
2	286	34.4	159	1	17KD_RICGJA
3	285	34.3	159	1	17KD_RICCN
4	276.5	33.2	159	1	17KD_RICTY
5	268	32.2	154	1	17KD_RICAU
6	259	31.1	154	1	17KD_RICPA
7	259	31.1	154	1	17KD_RICRH
8	256	30.8	154	1	17KD_RICMO
9	255	30.6	154	1	17KD_RICAM
10	156	18.8	80	1	17KD_RICCA
11	112.5	13.5	155	1	PCP_YREN
12	105.5	12.7	155	1	SLYB_SALTY
13	104.5	12.6	155	1	SLYB_ECOLI
14	101	12.1	155	1	PCP_HABIN
15	100	12.0	179	1	YCFJ_ECOLI
16	95	11.4	172	1	YFGH_ECOLI
17	90	10.8	526	1	K1CJ_BOVIN
18	90	10.8	526	1	K1CJ_BOVIN
19	88	10.6	1585	1	YQBO_BACSU
20	87.5	10.5	783	1	YKR2_CAEEL
21	86.5	10.4	72	1	OSMB_SALTY
22	85.5	10.3	72	1	OSMB_ECOLI
23	84.5	10.2	431	1	KR22_CANAL
24	82.5	9.9	243	1	CYSH_SALTY
25	82.5	9.9	541	1	NUS7_YEAST
26	82	9.9	132	1	Y615_AQUAE
27	81.5	9.8	243	1	CYSH_SALTY
28	81.5	9.8	301	1	STYG_RAT
29	81.5	9.8	526	1	VP5_BTUV1
30	80	9.6	806	1	ITB7_MOUSE
31	79.5	9.6	263	1	CANS_BOVIN
32	79.5	9.6	890	1	SYA_STRCO
33	79	9.5	593	1	K1CJ_HUMAN

34 78.5 9.4 219 1 YIAD_ECOLI
35 78.5 9.4 747 1 SPD1_NEPCL
36 78 9.4 514 1 ATPA_THIPE
37 77.5 9.3 359 1 ATPA_BOVIN
38 77.5 9.3 467 1 HEMI_MYCLE
39 77.5 9.3 543 1 ATPA_RAT
40 77.5 9.3 553 1 ATPA_HUMAN
41 77 9.3 266 1 CANS_RABIT
42 76.5 9.2 553 1 ATP0_BOVIN
43 76.5 9.2 553 1 ATPA_MOUSE
44 76.5 9.2 569 1 K1CJ_MOUSE
45 76.5 9.2 727 1 IF2M_HUMAN

P37665 escherichia
P19837 nephila cla
P41167 thiobacillu
P19482 bos taurus
P46724 mycobacteri
P15999 rattus norv
P25705 homo sapien
P06813 oryctolagus
Q03265 mus taurus
P02535 mus musculu
P46199 homo sapien

ALIGNMENTS

RESULT 1
17KD_RICPR
ID 17KD_RICPR STANDARD; PRT; 159 AA.
AC P16624;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 17 kDa surface antigen precursor.
GN OMP OR RP833.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiaceae; Rickettsia.
OX NCBI_TaxID=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Madrid E;
EX MEDLINE=89359171; PubMed=2768201;
RA Anderson B.E., Tzianabos T.;
RT "Comparative sequence analysis of a genus-common rickettsial antigen gene."
RL J. Bacteriol. 171:5199-5201(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Madrid E;
RX MEDLINE=99039499; PubMed=9823893;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA Sacheritz-Ponten T., Alemark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of mitochondria."
RL Nature 396:133-140(1998).
CC -!- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid anchor (Probable).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC EMBL; M28482; AAA26378.1; ALT_SEQ.
CC EMBL; AJ235273; CAA15258.1; -.
CC PIR; D33971; D33971.
CC PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Outer membrane; lipoprotein; Antigen; Signal; Complete proteome.
FT SIGNAL 1 19
FT CHAIN 20 159 17 kDa SURFACE ANTIGEN.
FT LIPID 20 20 N-ACYL DIGLYCERIDE (PROBABLE).
SQ SEQUENCE 159 AA; 16672 MW; A33D404B65EEB071 CRC64;

Query Match 34.6%; Score 288; DB 1; Length 159;
Best Local Similarity 37.9%; Pred. No. 3.9e-18;
Matches 61; Conservative 32; Mismatches 52; Indels 16; Gaps 6;

RC STRAIN=Maculatum;
RA Pretzman C.I., Stothard D.R., Ralph D.A., Clark J.B., Fuerst P.A.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
CC anchor (Probable).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U17008; AAB82040.1; -
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Outer membrane; Lipoprotein; Antigen; Signal.
FT SIGNAL 1 19 BY SIMILARITY.
FT CHAIN 20 >154 17 kDa SURFACE ANTIGEN.
FT LIPID 20 20 N-ACYL DIGLYCERIDE (PROBABLE).
FT NON_TER 154 154
SQ SEQUENCE 154 AA; 15897 MW; 5D06F45F9DBD5EEC CRC64;

Query Match 31.1%; Score 259; DB 1; Length 154;
Best Local Similarity 36.5%; Pred. No. 1.2e-15;
Matches 57; Conservative 30; Mismatches 53; Indels 16; Gaps 6;

QY 8 SLLIIISV---FLVGC--AQNFSROEVGAATGAVVGVAGQLFGKSGRVSMAG-GAVL 61
DB 5 SKIMVIALATSMQLACNGPGMNKQGTTLGGAGALLGSGFGKKGQL-VGVGVGALL 63
QY 62 GGLIGSKIGSQMDQDK----IKLNQSLKVKAGQVTRWRNPDTGNSYSVEPRTYQRYN 117
DB 64 GAVLGGQIGAGMDEQDRRLAELTSQRALETAPSGSNVWENPDNGNGYVTENKTY---- 119

RESULT 7
17KD_RICRH STANDARD; PRT; 154 AA.
ID 17KD_RICRH
AC P50931;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE 17 kDa surface antigen precursor (Fragment).
GN OMP.
OS Rickettsia rhipicephali.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsieae; Rickettsia.
OX NCBI_TaxID=33992;
RN [1]
RP SEQUENCE FROM N.A.
RC Stothard D.R., Ralph D.A., Clark J.B., Fuerst P.A., Pretzman C.;
RA Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
CC anchor (Probable).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U11020; AAB07706.1; -
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Outer membrane; Lipoprotein; Antigen; Signal.
FT SIGNAL 1 19 BY SIMILARITY.
FT CHAIN 20 >154 17 kDa SURFACE ANTIGEN.
FT LIPID 20 20 N-ACYL DIGLYCERIDE (PROBABLE).
FT NON_TER 154 154
SQ SEQUENCE 154 AA; 15897 MW; 5D06F45F9DBD5EEC CRC64;

Query Match 31.1%; Score 259; DB 1; Length 154;
Best Local Similarity 36.5%; Pred. No. 1.2e-15;
Matches 57; Conservative 30; Mismatches 53; Indels 16; Gaps 6;

QY 8 SLLIIISV---FLVGC--AQNFSROEVGAATGAVVGVAGQLFGKSGRVSMAG-GAVL 61
DB 5 SKIMVIALATSMQLACNGPGMNKQGTTLGGAGALLGSGFGKKGQL-VGVGVGALL 63
QY 62 GGLIGSKIGSQMDQDK----IKLNQSLKVKAGQVTRWRNPDTGNSYSVEPRTYQRYN 117
DB 64 GAVLGGQIGAGMDEQDRRLAELTSQRALETAPSGSNVWENPDNGNGYVTENKTY---- 119

RESULT 7
17KD_RICRH STANDARD; PRT; 154 AA.
ID 17KD_RICRH
AC P50931;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE 17 kDa surface antigen precursor (Fragment).
GN OMP.
OS Rickettsia rhipicephali.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsieae; Rickettsia.
OX NCBI_TaxID=33992;
RN [1]
RP SEQUENCE FROM N.A.
RC Stothard D.R., Ralph D.A., Clark J.B., Fuerst P.A., Pretzman C.;
RA Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
CC anchor (Probable).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U11020; AAB07706.1; -
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Outer membrane; Lipoprotein; Antigen; Signal.
FT SIGNAL 1 19 BY SIMILARITY.
FT CHAIN 20 >154 17 kDa SURFACE ANTIGEN.

FT LIPID 20 20 N-ACYL DIGLYCERIDE (PROBABLE).
FT NON_TER 154 154
SQ SEQUENCE 154 AA; 15895 MW; 0CF85AD5D96DFEFB CRC64;

Query Match 31.1%; Score 259; DB 1; Length 154;
Best Local Similarity 36.5%; Pred. No. 1.2e-15;
Matches 57; Conservative 30; Mismatches 53; Indels 16; Gaps 6;

QY 8 SLLIIISV---FLVGC--AQNFSROEVGAATGAVVGVAGQLFGKSGRVSMAG-GAVL 61
DB 5 SKIMVIALATSMQLACNGPGMNKQGTTLGGAGALLGSGFGKKGQL-VGVGVGALL 63
QY 62 GGLIGSKIGSQMDQDK----IKLNQSLKVKAGQVTRWRNPDTGNSYSVEPRTYQRYN 117
DB 64 GAVLGGQIGAGMDEQDRRLAELTSQRALETAPSGSNVWENPDNGNGYVTENKTY---- 119

RESULT 8
17KD_RICMO STANDARD; PRT; 154 AA.
ID 17KD_RICMO
AC P50929;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE 17 kDa surface antigen precursor (Fragment).
GN OMP.
OS Rickettsia montana.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsieae; Rickettsia.
OX NCBI_TaxID=33991;
RN [1]
RP SEQUENCE FROM N.A.
RC Stothard D.R., Ralph D.A., Clark J.B., Fuerst P.A., Pretzman C.;
RA Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
CC anchor (Probable).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U11017; AAB07705.1; -
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Outer membrane; Lipoprotein; Antigen; Signal.
FT SIGNAL 1 19 BY SIMILARITY.
FT CHAIN 20 >154 17 kDa SURFACE ANTIGEN.
FT LIPID 20 20 N-ACYL DIGLYCERIDE (PROBABLE).
FT NON_TER 154 154
SQ SEQUENCE 154 AA; 15881 MW; A09C53B8769E31DA CRC64;

Query Match 30.8%; Score 256; DB 1; Length 154;
Best Local Similarity 36.5%; Pred. No. 2.2e-15;
Matches 57; Conservative 30; Mismatches 53; Indels 16; Gaps 6;

QY 8 SLLIIISV---FLVGC--AQNFSROEVGAATGAVVGVAGQLFGKSGRVSMAG-GAVL 61
DB 5 SKIMVIALATSMQLACNGPGMNKQGTTLGGAGALLGSGFGKKGQL-VGVGVGALL 63
QY 62 GGLIGSKIGSQMDQDK----IKLNQSLKVKAGQVTRWRNPDTGNSYSVEPRTYQRYN 117
DB 64 GAVLGGQIGAGMDEQDRRLAELTSQRALETAPSGSNVWENPDNGNGYVTENKTY---- 119

RESULT 8
17KD_RICMO STANDARD; PRT; 154 AA.
ID 17KD_RICMO
AC P50929;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE 17 kDa surface antigen precursor (Fragment).
GN OMP.
OS Rickettsia montana.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsieae; Rickettsia.
OX NCBI_TaxID=33991;
RN [1]
RP SEQUENCE FROM N.A.
RC Stothard D.R., Ralph D.A., Clark J.B., Fuerst P.A., Pretzman C.;
RA Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
CC anchor (Probable).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U11017; AAB07705.1; -
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Outer membrane; Lipoprotein; Antigen; Signal.
FT SIGNAL 1 19 BY SIMILARITY.
FT CHAIN 20 >154 17 kDa SURFACE ANTIGEN.
FT LIPID 20 20 N-ACYL DIGLYCERIDE (PROBABLE).
FT NON_TER 154 154
SQ SEQUENCE 154 AA; 15881 MW; A09C53B8769E31DA CRC64;

Query Match 30.8%; Score 256; DB 1; Length 154;
Best Local Similarity 36.5%; Pred. No. 2.2e-15;
Matches 57; Conservative 30; Mismatches 53; Indels 16; Gaps 6;

QY 8 SLLIIISV---FLVGC--AQNFSROEVGAATGAVVGVAGQLFGKSGRVSMAG-GAVL 61
DB 5 SKIMVIALATSMQLACNGPGMNKQGTTLGGAGALLGSGFGKKGQL-VGVGVGALL 63
QY 62 GGLIGSKIGSQMDQDK----IKLNQSLKVKAGQVTRWRNPDTGNSYSVEPRTYQRYN 117
DB 64 GAVLGGQIGAGMDEQDRRLAELTSQRALETAPSGSNVWENPDNGNGYVTENKTY---- 119

RESULT 8
17KD_RICMO STANDARD; PRT; 154 AA.
ID 17KD_RICMO
AC P50929;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE 17 kDa surface antigen precursor (Fragment).
GN OMP.
OS Rickettsia montana.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsieae; Rickettsia.
OX NCBI_TaxID=33991;
RN [1]
RP SEQUENCE FROM N.A.
RC Stothard D.R., Ralph D.A., Clark J.B., Fuerst P.A., Pretzman C.;
RA Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
CC anchor (Probable).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U11017; AAB07705.1; -
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Outer membrane; Lipoprotein; Antigen; Signal.
FT SIGNAL 1 19 BY SIMILARITY.
FT CHAIN 20 >154 17 kDa SURFACE ANTIGEN.

```
Db 120 -RNSTGQYCREYQTQTVVIGGKQKQAYGNACLPDQG 154
RESULT 9
17KD_RICAM
ID 17KD_RICAM STANDARD; PRT; 154 AA.
AC P50927;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE 17 kDa surface antigen precursor (Fragment).
GN OMP.
OS Rickettsia amblyommi.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=33989;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MO 95-1084;
RA Stothard D.R., Ralph D.A., Clark J.B., Fuerst P.A., Pretzman C.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
CC anchor (Probable).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M82879; -; NOT ANNOTATED CDS.
DR InterPro; IPR000437; Prok_lipoProt.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; PARTIAL.
KW Outer membrane; Lipoprotein; Antigen.
FT NON_TER 1
FT NON_TER 80
SQ SEQUENCE 80 AA; 8372 MW; AD289A48EAB19E08 CRC64;
Query Match 18.8%; Score 156; DB 1; Length 80;
Best Local Similarity 40.3%; Pred. No. 4.8e-07;
Matches 31; Conservative 15; Mismatches 27; Indels 4; Gaps 1;
QY 42 GOLFQKSGRVSMAIGAVLGGLGKIGKIQSDQDDK----IKLNQSLKVKAGQVTRWR 97
Db 1 GSQFCKGKGLIGVCGALLGAILGNQICAGWDEQDRRLAELTSQRALETTSGTSIEWR 60
QY 98 NPDTCNSYSVEPRTYQ 114
Db 61 NPDNGNGYVTPSKTYK 77
RESULT 11
PCP_YEREN
ID PCP_YEREN STANDARD; PRT; 155 AA.
AC P31484;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Outer membrane lipoprotein pcp precursor.
GN PCP OR PCPY.
OS Yersinia enterocolitica.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=630;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 51872 / WA-C / Serotype O:8;
RX MEDLINE=92121089; PubMed=1732192;
RA Baemler A.J., Hantke K.;
RT "A lipoprotein of Yersinia enterocolitica facilitates ferrioxamine
RT uptake in Escherichia coli.";
RL J. Bacteriol. 174:1029-1035(1992).
CC -!- FUNCTION: FACILITATES FERRIOXAMINE UPTAKE.
CC -!- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
CC anchor.
CC -!- SIMILARITY: TO E.COLI AND S.TYPHIMURIUM SLI5B AND TO
CC H.INFLUENZAE PCP.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
Db 120 -RNSTGQYCREYQTQTVVIGGKQKQAYGNACLPDQG 154
RESULT 10
17KD_RICCA
ID 17KD_RICCA STANDARD; PRT; 80 AA.
AC P29697;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE 17 kDa surface antigen (Fragment).
GN OMP.
OS Rickettsia canada.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=788;
RN [1]
RP SEQUENCE FROM N.A.
```

```

DR EMBL; X60448; CAA42977.1; -.
DR PIR; S23787; S23787.
DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.
KW Outer membrane; Lipoprotein; Signal.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 155 OUTER MEMBRANE LIPOPROTEIN PCP.
FT LIPID 18 18 N-ACYL DIGLYCERIDE (POTENTIAL).
SQ SEQUENCE 155 AA; 15362 MW; 8AD6BE2132E849FA CRC64;

Query Match 13.5%; Score 112.5; DB 1; Length 155;
Best Local Similarity 24.7%; Pred. No. 0.0053;
Matches 37; Conservative 19; Mismatches 43; Indels 51; Gaps 4;

QY 9 SLIIISVFLVGCQN-----FSROE-----VGAAT 33
DB 7 AVAAATVLTGCANNLTSGDVSQAQKVQTVTGTLLSVRPTVITQGGDDNNVGAIG 66

QY 34 GAVGVGVAGOLFCKGSRVSMATGGAVLGLGSKIGSQMDQDKL-----81
DB 67 GAVLGGFLGNTVGGTGRSLATAGAVAGMGAGQGVQGMNRTDGVQLEVRKDDGTTILV 126

QY 82 --NQSEKVKAGQVTRWNPDTGNSVSEP 109
DB 127 VQKQPTREFSVGO--RVMLASSGSTVTVSP 154

RESULT 12
SLYB_SALTY STANDARD; PRT; 155 AA.
AC Q53549;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Outer membrane lipoprotein slyb precursor.
GN SLYB OR STM1445 OR STV1677 OR TI313.
OS Salmonella typhimurium, and
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602, 601;

[1]
SEQUENCE FROM N.A.
SPECIES=S.typhimurium;
MEDLINE=21534948; PubMed=11677609;
Ludwig A., Tengel C., Bauer S., Hubert A., Benz R., Mollenkopf H.-J.,
Goebel W.;
"SLyA, a regulatory protein from Salmonella typhimurium, induces a
haemolytic and pore-forming protein in Escherichia coli.";
Mol. Gen. Genet. 249:474-486(1995).
[2]
SEQUENCE FROM N.A.
SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
MEDLINE=21534948; PubMed=11677609;
McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
Waterston R., Wilson R.K.;
"Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2.";
Nature 413:852-856(2001).
[3]
SEQUENCE FROM N.A.
SPECIES=S.typhi; STRAIN=CT18;
MEDLINE=21534947; PubMed=11677608;
Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
Whitehead S., Barrell B.G.;

"Complete genome sequence of a multiple drug resistant Salmonella
enterica serovar Typhi CT18.";
Nature 413:848-852(2001).
[4]
SEQUENCE FROM N.A.
SPECIES=S.typhi; STRAIN=Ty2 / ATCC 700931;
MEDLINE=22531367; PubMed=12644504;
Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
"Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
and CT18.";
J. Bacteriol. 185:2330-2337(2003).
RL J. SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
anchor (Potential).
CC -!- SIMILARITY: TO E.COLI SLYB, H.INFLUENZAE PCP AND Y. ENTEROCOLITICA
PCP.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; S80790; AAB35871.2; -.
DR EMBL; AE008762; AAL20367.1; -.
DR EMBL; AL627271; CAD01922.1; -.
DR EMBL; AE016838; AAO88963.1; -.
DR StyGene; SG10573; slyb.
DE Outer membrane; Lipoprotein; Signal; Complete proteome.
KW PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 155 OUTER MEMBRANE LIPOPROTEIN SLYB.
FT LIPID 18 18 N-ACYL DIGLYCERIDE.
SQ SEQUENCE 155 AA; 15548 MW; 82FDDCDBAB55A7 CRC64;

Query Match 12.7%; Score 105.5; DB 1; Length 155;
Best Local Similarity 24.3%; Pred. No. 0.021;
Matches 34; Conservative 20; Mismatches 53; Indels 33; Gaps 3;

QY 1 MRGCLQGSSL-----IIISVFLVGCQNFSRQEVCAATGAVVGVGA 41
DB 15 LAGCVNDSLGDVYTASEAKQVNVYGTIVNRPVQIQGGDDSNVIGAGVIGGFL 74

QY 42 GQLFKGGRSVSMATGGAVLGLGSKIGSQMDQDKIKLQSLKVKAGQVTRWNPDT 101
DB 75 GNTIGGTRSLATAGAVAGVAGVAGQVQSAANKTQGVLEL--EIRK-----DD 120

QY 102 GNSYSVEPVRTYQRYNKOER 121
DB 121 GNTIMVQKQNTREFSAGOR 140

RESULT 13
SLYB_ECOLI STANDARD; PRT; 155 AA.
AC P55741; P76183;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Outer membrane lipoprotein slyb precursor.
GN SLYB OR B1641 OR Z2655 OR ECS2350 OR SF1668.
OS Escherichia coli,
OS Escherichia coli O157:H7, and
OS Shigella flexneri
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 83334, 623;
[1]
SEQUENCE FROM N.A.
SPECIES=E.coli;
MEDLINE=96133688; PubMed=8544813;

```


"Whole-genome random sequencing and assembly of Haemophilus influenzae Rd."
RL Science 269:496-512(1995).
CC -!- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid anchor
CC
CC -!- SIMILARITY: TO E.COLI AND S.TYPHIMURIUM SLVB AND TO Y. ENTEROCOLITICA PCP.
CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC

DR EMBL; M18877; AAA24938.1; -;
DR EMBL; U32832; AAC23228.1; -;
DR PIR; I64130; I64130.
DR TIGR; H11579; -;
DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.
KW Outer membrane; Lipoprotein; Signal; Complete proteome.
FT SIGNAL 1 18
FT CHAIN 19 155 OUTER MEMBRANE LIPOPROTEIN PCP.
FT LIPID 19 19 N-ACYL DIGLYCERIDE.
FT CONFLICT 135 143 CSLVAEFVF -> VAGRRVRI (IN REF. 1).
SQ SEQUENCE 155 AA; 15425 MW; D7880327FCFC985 CRC64;

Query Match 12.1%; Score 101; DB 1; Length 155;
Best Local Similarity 39.8%; Pred. No. 0.052;
Matches 21; Conservative 9; Mismatches 23; Indels 0; Gaps 0;

QY 29 VGAATGAVGVGGVAGLFGKSGRVSMAGVGLGVLGSGKIGQSMDDQDKIKL 81
DB 62 VGTGLGALGGINGSTIGGRGQQAIAVVGATGATAGTKIEKMSQVNGAEL 114

RESULT 15

ID YCFJ ECOLI STANDARD; PRT; 179 AA.
AC P37756; P75951;
DT 01-OCT-1994 (Rel. 30, Created)
DT 28-FEB-2003 (Rel. 35, Last sequence update)
DE Hypothetical protein ycfJ.
GN YCFJ OR B1110 OR C1383 OR SF1114.
OS Escherichia coli, O6, and
OS Escherichia coli O6, and
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 217992, 623;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y.;
RA "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN

RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A., Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K., Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K., Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N., Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y., Yano M., Horiuchi T.;
RA "A 718-Kb DNA sequence of the Escherichia coli K-12 genome

RT corresponding to the 12.7-28.0 min region on the linkage map."
RL DNA Res. 3:137-155(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P., Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D., Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T., Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RA "Extensive mosaic structure revealed by the complete genome sequence of uropathogenic Escherichia coli."
RT Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
RL [4]
RN SEQUENCE OF 1-63 FROM N.A.
RC SPECIES=E.coli; STRAIN=K12;
RX MEDLINE=81236546; PubMed=6265208;
RA Young J.G., Rogers B.L., Campbell H.D., Jaworski A., Shaw D.C.;
RT "Nucleotide sequence coding for the respiratory NADH dehydrogenase of Escherichia coli. UUG initiation codon."
RL Eur. J. Biochem. 116:165-170(1981).
RN [5]
RP SEQUENCE FROM N.A.
RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H., Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J., Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S., Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y., Yu J.;
RA "Genome sequence of Shigella flexneri 2a: insights into pathogenicity through comparison with genomes of Escherichia coli K12 and O157."
RT Nucleic Acids Res. 30:4432-4441(2002).
RL [6]
RN IDENTIFICATION.
RP SPECIES=E.coli;
RX MEDLINE=95075659; PubMed=7984428;
RA Borodovsky M., Rudd K.E., Koonin E.V.;
RT "Intrinsic and extrinsic approaches for detecting genes in a bacterial genome."
RL Nucleic Acids Res. 22:4756-4767(1994).
CC -!- SIMILARITY: TO RICKETTSIA 17 KDA SURFACE ANTIGEN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000211; AAC74194.1; -;
DR EMBL; D90746; BAA35925.1; -;
DR EMBL; AE016759; AAN79853.1; ALT INIT.
DR EMBL; V00306; -; NOT ANNOTATED CDS.
DR EMBL; AE015138; AAN42732.1; ALT INIT.
DR PIR; C64855; C64855.
DR EcGene; BG12444; ycfJ.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 5 25 POTENTIAL.
SQ SEQUENCE 179 AA; 18920 MW; BA5EB0DB56D45609 CRC64;
Query Match 12.0%; Score 100; DB 1; Length 179;
Best Local Similarity 26.3%; Pred. No. 0.073;
Matches 36; Conservative 17; Mismatches 48; Indels 36; Gaps 4;
QY 30 GAATGAVGVGGVAGLFGKSGRVSMAGVGLGVLGSGKIGQSMDDQDKIKNQSLKVK 89
DB 73 GSVLGAVGAGVGHQFGGGRGKDVATVVGALGGVAGNQGSLQESD----- 120
QY 90 AGQVTRWRNPDTGNSYSEVPTVYQRNKRQRYCREFOQKAMIAQKQEIY-----G 144
DB 121 -----TYTTTQQRCKTVYDKSEKMLGVDVYTK-----IGDQGGKIRMDRDPG 162

QY 145 TACPQDGRWQVISTEK 161
| | | | |
Db 163 TQIPL-DSNGQLLNK 178

Search completed: November 5, 2003, 20:13:25
Job time : 6.39551 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 20:09:46 ; Search time 23.6356 Seconds
(without alignments)
1757.794 Million cell updates/sec

Title: US-09-677-374-4
Perfect score: 832
Sequence: 1 MRGCLQGSSLIISVFLVGC.....IVGTACPOPDGRWQVISTEK 161

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	815	98.0	162 2 Q9F9K8	Q9F9K8 pisciricket
2	303.5	36.5	148 2 O54381	O54381 rickettsia
3	283	34.0	159 2 Q9F9F2	Q9F9F2 rickettsia
4	258.5	31.1	137 2 O52252	O52252 rickettsia
5	252.5	30.3	144 2 Q9K2N6	Q9K2N6 male-killin
6	251.5	30.2	137 2 O31065	O31065 rickettsia
7	251.5	30.2	144 2 Q9K4W8	Q9K4W8 male-killin
8	244	29.3	154 2 O53154	O53154 rickettsia
9	241.5	29.0	131 2 Q8KL74	Q8KL74 rickettsia
10	241.5	29.0	131 2 Q8KL75	Q8KL75 rickettsia
11	239	28.7	151 2 Q9F9Q9	Q9F9Q9 rickettsia
12	236.5	28.4	131 2 Q9F0Q1	Q9F0Q1 rickettsia
13	236.5	28.4	131 2 Q52637	Q52637 rickettsia
14	235.5	28.3	131 2 Q8KLU2	Q8KLU2 rickettsia
15	235.5	28.3	131 2 Q91522	Q91522 rickettsia
16	216.5	26.0	105 2 O31208	O31208 rickettsia

ALIGNMENTS

RESULT 1

Q9F9K8 PRELIMINARY; PRT; 162 AA.
ID AC Q9F9K8;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE 17 kDa antigen.
GN OSPR.
OS Piscirickettsia salmonis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Thiotrichales;
OC Piscirickettsiaceae; Piscirickettsia.
OX NCBI_TaxID=1238;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LF-89;
RA Kuzk M.A., Burian J., Thornton J.C., Kay W.W.;
RT "Identification of a genus-common Rickettsial surface antigen in the salmonid pathogen Piscirickettsia salmonis.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF184152; AAG17000.1; -.
SQ SEQUENCE 162 AA; 17661 MW; DDE99B6FD94A527E CRC64;

Query Match 98.0%; Score 815; DB 2; Length 162;
Best Local Similarity 98.8%; Pred. No. 1.3e-65;
Matches 158; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 2 RCLQSSLIISVFLVGCANFNRQEVGAATGAVVGVGAGQLFGKSGRVSMAIGAVL 61
Db 3 RCLQSSLIISVFLVGCANFNRQEVGAATGAVVGVGAGQLFGKSGRVSMAIGAVL 62
QY 62 GGLGSKIGSQSDQDDPKIKLNQSLKVKAGQVTRWRNPDGTGNSYSVEPVRTYQRYNKQR 121
Db 63 GGLGSKIGSQSDQDDPKIKLNQSLKVKAGQVTRWRNPDGTGNSYSVEPVRTYQRYNKQR 122
QY 122 RQYCRFFQOKAMIAQOQEIYGTACPDGRWQVISTEK 161
Db 123 RQYCRFFQOKAMIAQOQEIYGTACPDGRWQVISTEK 162

Q9AGC7 rickettsia
Q985G4 rhizobium l
Q9GLU7 brucella su
Q9HXI3 pseudomonas
Q8VUE8 brucella ab
Q8YGT27 brucella me
Q92R89 rhizobium m
Q914S1 pseudomonas
Q8GEL6 pseudomonas
Q8KCR0 chlorobium
Q8U5V9 agrobacteri
Q8TMM1 methanosarc
Q8XHL3 salmonella
Q9A3X8 caulobacter
Q9PGX0 xylella fas
Q9F6B1 edwardsiell
Q928T9 rhizobium m
Q8UGR3 agrobacteri
Q8ZEL6 yersinia pe
Q9RA95 serratia sp
Q9PAQ1 xanthomonas
Q9RB08 pectobacter
Q52854 rhizobium l
Q9P5A0 xanthomonas
Q8DIN7 yersinia pe
Q8G099 brucella su
Q8YHL1 brucella me
Q9XCA4 porphyromon
Q8CWL1 escherichia

```

RESULT 2
OS4381 ID O54381 PRELIMINARY; PRT; 148 AA.
AC O54381;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE 17 kDa common-antigen (Fragment).
OS Rickettsia sp.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=789;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98087556; PubMed=9425244;
RA Davis M.J., Ying Z., Brunner B.R., Pantoja A., Ferwerda F.H.;
RT "Rickettsial relative associated with papaya bunchy top disease.";
RL Curr. Microbiol. 36:80-84(1998).
DR EMBL; U76907; AAC02809.1; -.
FT NON_TER 1
FT NON_TER 148
SQ SEQUENCE 148 AA; 15050 MW; A7AFEEFDE0AEEB4C CRC64;

Query Match 36.5%; Score 303.5; DB 2; Length 148;
Best Local Similarity 40.1%; Pred. No. 1.1e-19;
Matches 55; Conservative 30; Mismatches 43; Indels 9; Gaps 2;

QY 25 SROEVAATGAVVGGVAGQLFGKSGRVSMAIGAVLGLGSKIGQSDQDK-----IK 80
Db 17 NKQSGTLLGTLGVLGSGFGGSGTGLAAVAGAGALLGAILGNQIGAGHDEQDRKLAELT 76

QY 81 LNOSLEKVKAGQVTRWRNPDTGNSYVPEVRYQYKQERRQYCRFQOKAMIAQOK 140
Db 77 SQRALEAAPSQSSVWRNPNGNYGVTFSKAY-----KNTGQYCREYTTQVTVVGKQ 131

QY 141 EYGTACPDGRQVQVI 157
Db 132 KAYGTACRQPDGQVQV 148

RESULT 3
Q9F9F2 ID Q9F9F2 PRELIMINARY; PRT; 159 AA.
AC Q9F9F2;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE 17 kDa genus-common antigen.
OS Rickettsia felis (Rickettsia azadi).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=42862;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21217364; PubMed=11321078;
RA Bouyer D.H., Stenos J., Crocquet-Valdes P., Moron C.G., Popov V.L.,
RA Zavala-Velazquez J.E., Foil L.D., Stothard D.R., Azad A.F.,
RA Walker D.H.;
RT "Rickettsia felis: molecular characterization of a new member of the
RT spotted fever group.";
RL Int. J. Syst. Evol. Microbiol. 51:339-347(2001).
DR EMBL; AF195118; AAC28452.1; -.
SQ SEQUENCE 159 AA; 16497 MW; 34C5B020AF470A1F CRC64;

Query Match 34.0%; Score 283; DB 2; Length 159;
Best Local Similarity 37.9%; Pred. No. 8.6e-18;
Matches 61; Conservative 30; Mismatches 54; Indels 16; Gaps 6;

QY 8 SSLIIISV---FLVGC--AQNFSRQEVGAATGAVVGGVAGQLFGKSGRVSMAIG-CAVL 61
Db 5 SKIMIIAASMLQACNGPGGMKNQGTLLGGAGGALLGSQFGKGGQL-VGVGVGALL 63

```

```

QY 62 GGLGSKIGQSDQDK-----IKNQSLKVKAGQVTRWRNPDTGNSYVPEVRYQRYN 117
Db 64 GAVLGGQIGAGMDEQDRRLAELTSQRALEAPSGTSVSEWRNPNGNHGYVTPNKTY---- 119

QY 118 QERRRQYCRFQOKAMIAQOKQBIYGTACPDGRQVQVIS 158
Db 120 -RNSTGQYCREYTTQVTVGGKQKAYGNACRQPDGLWQVNV 159

RESULT 4
OS2252 ID O52252 PRELIMINARY; PRT; 137 AA.
AC O52252;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE 17 kDa antigen (Fragment).
OS Rickettsia cooley.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=69410;
RN [1]
RP SEQUENCE FROM N.A.
RX Billings A.N., Teltow G.J., Walker D.H.;
RT "Molecular characterization of a novel spotted fever group rickettsial
RT species from Ixodes scapularis in Texas.";
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF01534; AAB95267.1; -.
FT NON_TER 1
FT NON_TER 137
SQ SEQUENCE 137 AA; 14215 MW; A27597A9AFD85FC3 CRC64;

Query Match 31.1%; Score 258.5; DB 2; Length 137;
Best Local Similarity 39.3%; Pred. No. 1.1e-15;
Matches 53; Conservative 25; Mismatches 46; Indels 11; Gaps 4;

QY 22 QNFSRQEVGAATGAVVGGVAGQLFGKSGRVSMAIG-CAVLGLGSKIGQSDQDK-- 78
Db 7 RGMNKQGTGTLGGAGGALLGSQFGKGGQL-VGVGVGALLGAVLGGQIGAGMDEQDRRL 65

QY 79 --IKNQSLKVKAGQVTRWRNPDTGNSYVPEVRYQYKQERRQYCRFQOKAMIA 136
Db 66 BELTSQRALEAAPSQSSVWRNPNGNYGVTPNKTY-----RNSTGQYCREYTTQVTVIG 120

QY 137 GQKQBIYGTACPD 151
Db 121 GQKQKAYGNACRQPD 135

RESULT 5
Q9K2N6 ID Q9K2N6 PRELIMINARY; PRT; 144 AA.
AC Q9K2N6;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE 17kDa antigen (17 kDa antigen) (Fragment).
OS male-killing Rickettsia from Adalia bipunctata.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=38028;
RN [1]
RP SEQUENCE FROM N.A.
RX Schlenker H.J.G.V.D., Habig M., Sloggett J.J., Webberley M.K.,
RA Bertrand D., Hurst G.D., Majerus M.E.N.;
RT "On the evolution of male-killing: Monophyletic origin and horizontal
RT transfer of male-killing Rickettsia (a-Proteobacteria) from two con-
RT generic ladybirds, Adalia bipunctata L. and A. decempunctata L.
RT (Coleoptera: Coccinellidae).";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ269518; CAB96383.1; -.
DR EMBL; AJ269517; CAB96382.1; -.
FT NON_TER 1
FT NON_TER 144

```

SQ SEQUENCE 144 AA; 14785 MW; C8254739CCAS6AE7 CRC64;
 Query Match 30.3%; Score 252.5; DB 2; Length 144;
 Best Local Similarity 39.4%; Pred. No. 4.2e-15;
 Matches 52; Conservative 24; Mismatches 45; Indels 11; Gaps 4;
 QY 25 SRQEVGAATGAVGVAGQLFGKSGRVSMIAIG-GAVLGLIGSKIGSMDOODK----I 79
 Db 17 NKQGTGTLGGAGGALLGSQFGKGGQL-VGVGVGALLGAVLGGIGAGMDQDRRLAEL 75
 QY 80 KLNQSLKVKAGQVTRWNPDTGNSYSVEPVRTYQRYNKOERRQYCRFQOKAMIAGOK 139
 Db 76 TSQRALEAAPSGNSVWERNPDNGNHYVTPNKTY-----RNSTGQYCREYTTVVIGGQK 130
 QY 140 QEYGTACPPD 151
 Db 131 QKAYGNACRPD 142
 RESULT 6
 ID O31065 PRELIMINARY; PRT; 137 AA.
 AC O31065; Q9HW02;
 DT 01-JAN-1998 (TRENBLrel. 05, Created)
 DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
 DE 17 kDa antigen (17 kDa protein) (Fragment).
 OS Rickettsia honei.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsiidae; Rickettsia.
 OX NCBI_TaxID=37816;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TT-118;
 RA Billings A.N., Yu X.-J., Teel P.D., Walker D.H.;
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99045882; PubMed=9828442;
 RA Stenos J., Roux V., Walker D., Raoult D.;
 RT "Rickettsia honei sp. nov., the aetiological agent of Flinders Island
 spotted fever in Australia."
 RL Int. J. Syst. Bacteriol. 48:1399-1404 (1998).
 DR EMBL; AF027124; AAB81846.1; -;
 DR EMBL; AF060706; AAD20231.1; -;
 DR EMBL; AF060704; AAD20230.1; -;
 FT NON_TER 1
 SQ SEQUENCE 137 AA; 14167 MW; 75BC1D0D745B428C CRC64;
 Query Match 30.2%; Score 251.5; DB 2; Length 137;
 Best Local Similarity 39.4%; Pred. No. 4.9e-15;
 Matches 52; Conservative 24; Mismatches 45; Indels 11; Gaps 4;
 QY 25 SRQEVGAATGAVGVAGQLFGKSGRVSMIAIG-GAVLGLIGSKIGSMDOODK----I 79
 Db 10 NKQGTGTLGGAGGALLGSQFGKGGQL-VGVGVGALLGAVLGGIGAGMDQDRRLAEL 68
 QY 80 KLNQSLKVKAGQVTRWNPDTGNSYSVEPVRTYQRYNKOERRQYCRFQOKAMIAGOK 139
 Db 69 TSQRALEAAPSGNSVWERNPDNGNHYVTPNKTY-----RNSTGQYCREYTTVVIGGQK 123
 QY 140 QEYGTACPPD 151
 Db 124 QKAYGNACRPD 135
 RESULT 7
 ID Q9K4W8 PRELIMINARY; PRT; 144 AA.
 AC Q9K4W8;
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)

DE 17 kDa antigen (Fragment).
 OS male-killing Rickettsia from Adalia decempunctata.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsiidae; Rickettsia.
 OX NCBI_TaxID=120393;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20575219; PubMed=11133455;
 RA Schulenburg H.J.G.V.D., Habis M., Sloggett J.J., Webberley M.K.,
 Bertrand D., Hurst G.D.D., Majerus M.E.N.;
 RT "Incidence of male-killing Rickettsia spp. (alpha-Proteobacteria) in
 the Ten-Spot Ladybird beetle Adalia decempunctata L. (Coleoptera:
 Coccinellidae)."
 RL Appl. Environ. Microbiol. 67:270-277 (2001).
 DR EMBL; AJ269516; CAB96381.1; -;
 FT NON_TER 1
 SQ SEQUENCE 144 AA; 14801 MW; C825472F16A56AE7 CRC64;
 Query Match 30.2%; Score 251.5; DB 2; Length 144;
 Best Local Similarity 39.4%; Pred. No. 5.2e-15;
 Matches 52; Conservative 24; Mismatches 45; Indels 11; Gaps 4;
 QY 25 SRQEVGAATGAVGVAGQLFGKSGRVSMIAIG-GAVLGLIGSKIGSMDOODK----I 79
 Db 17 NKQGTGTLGGAGGALLGSQFGKGGQL-VGVGVGALLGAVLGGIGAGMDQDRRLAEL 75
 QY 80 KLNQSLKVKAGQVTRWNPDTGNSYSVEPVRTYQRYNKOERRQYCRFQOKAMIAGOK 139
 Db 76 TSQRALEAAPSGNSVWERNPDNGNHYVTPNKTY-----RNSTGQYCREYTTVVIGGQK 130
 QY 140 QEYGTACPPD 151
 Db 131 QKAYGNACRPD 142
 RESULT 8
 ID Q53154 PRELIMINARY; PRT; 154 AA.
 AC Q53154;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
 DE (Clone PRB F15F 1), 5' end CDS (Fragment).
 OS Rickettsia sp.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsiidae; Rickettsia.
 OX NCBI_TaxID=789;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93084757; PubMed=1452660;
 RA Baird R.W., Lloyd M., Stenos J., Ross B.C., Stewart R.S., Dwyer B.;
 RT "Characterization and comparison of Australian human spotted fever
 group rickettsiae."
 RL J. Clin. Microbiol. 30:2896-2902 (1992).
 DR EMBL; M93931; AAA73386.1; -;
 FT NON_TER 154
 SQ SEQUENCE 154 AA; 15849 MW; F5C35855E8B439D2 CRC64;
 Query Match 29.3%; Score 244; DB 2; Length 154;
 Best Local Similarity 35.3%; Pred. No. 2.7e-14;
 Matches 55; Conservative 31; Mismatches 54; Indels 16; Gaps 6;
 QY 8 SSLIIISV---FLVGC--AQNFSRQEVGAATGAVGVAGQLFGKSGRVSMIAIG-GAVL 61
 Db 5 SKIMIIATATMLQACNCPGNGNKGTTGLGGAGGALLGSQFGKGGQL-VGVGVGALL 63
 QY 62 GGLIGSKIGSMDOODK----IKLNQSLKVKAGQVTRWNPDTGNSYSVEPVRTYQRYN 117
 Db 64 GAVLGGIGAGMDQDRRLAELTSQRALEAAPSGNSVWERNPDNGSYGVTPNKTYRST 123
 QY 118 KQERRQYCRFQOKAMIAGOKQEYGTACPPDGR 153
 Db 118 KQERRQYCRFQOKAMIAGOKQEYGTACPPDGR 153

Db 124 GOD-----CRVYQTQTVVIGGKQKAYGNACRQPDGQ 154

RESULT 9

Q8KLT4 PRELIMINARY; PRT; 131 AA.
ID Q8KLT4;
AC Q8KLT4;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Outer membrane protein (Fragment).
GN 17 KDA.
OS Rickettsia sp. IritA3.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=184232;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=IritA3;
RX MEDLINE=22182650; PubMed=12194779;
RA Beninati T., Lo N.L., Noda H., Esposito F., Rizzoli A., Favia G.,
RA Genchi C.;
RT "First detection of spotted fever group rickettsiae in Ixodes ricinus
from Italy.";
RL Emerg. Infect. Dis. 8:983-986(2002).
DR EMBL; AJ427883; CAD20879.1; -.
FT NON_TER 1
FT NON_TER 131
SQ SEQUENCE 131 AA; 13358 MW; D4152713C9FAA9CA CRC64;

Query Match 29.0%; Score 241.5; DB 2; Length 131;
Best Local Similarity 39.1%; Pred. No. 3.7e-14;
Matches 50; Conservative 24; Mismatches 43; Indels 11; Gaps 4;
QY 25 SRQEVGAATGAVVGVAGQFGKSGRVSMAIG-GAVLGGLIGSKIGQSMDOQDK----I 79
Db 10 NKQGTGTLGGAGGALLGQFGKGGQL-VGVGVGALLGAVLGGQIGAGMDEQDRRLAEL 68
QY 80 KLNOSLEKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKKERRQYCRFQOKAMTAGOK 139
Db 69 TSQRALEAAPSGSSTWRPNPDNGNYGYVTPNKTY-----RNSTGQYCREYTTQTVVIGGKQ 123
QY 140 QBIYGTAC 147
Db 124 QKAYGNAC 131

RESULT 10

Q8KLT5 PRELIMINARY; PRT; 131 AA.
ID Q8KLT5;
AC Q8KLT5;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Outer membrane protein (Fragment).
GN 17 KDA.
OS Rickettsia sp. IritA2.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=184231;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=IritA2;
RX MEDLINE=22182650; PubMed=12194779;
RA Beninati T., Lo N.L., Noda H., Esposito F., Rizzoli A., Favia G.,
RA Genchi C.;
RT "First detection of spotted fever group rickettsiae in Ixodes ricinus
from Italy.";
RL Emerg. Infect. Dis. 8:983-986(2002).
DR EMBL; AJ427882; CAD20878.1; -.
FT NON_TER 1
FT NON_TER 131
SQ SEQUENCE 131 AA; 13358 MW; D4152713C9FAA9CA CRC64;

Query Match 29.0%; Score 241.5; DB 2; Length 131;
Best Local Similarity 39.1%; Pred. No. 3.7e-14;
Matches 50; Conservative 24; Mismatches 43; Indels 11; Gaps 4;

QY 25 SRQEVGAATGAVVGVAGQFGKSGRVSMAIG-GAVLGGLIGSKIGQSMDOQDK----I 79
Db 10 NKQGTGTLGGAGGALLGQFGKGGQL-VGVGVGALLGAVLGGQIGAGMDEQDRRLAEL 68
QY 80 KLNOSLEKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKKERRQYCRFQOKAMTAGOK 139
Db 69 TSQRALEAAPSGSSTWRPNPDNGNYGYVTPNKTY-----RNSTGQYCREYTTQTVVIGGKQ 123
QY 140 QBIYGTAC 147
Db 124 QKAYGNAC 131

RESULT 11

Q3F9Q9 PRELIMINARY; PRT; 151 AA.
ID Q3F9Q9;
AC Q3F9Q9;
DT 01-MAR-2001 (T-EMBLrel. 16, Created)
DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (T-EMBLrel. 22, Last annotation update)
DE Outer membrane protein (Fragment).
OS Rickettsia helvetica.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=35789;
RN [1]
RP SEQUENCE FROM N.A.
RA Nilsson K., Pahlson C.;
RT "Novel peptide diagnostic reagent and kit for detection of
rickettsiosis.";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF181036; AAG09427.1; -.
FT NON_TER 151
SQ SEQUENCE 151 AA; 15621 MW; B77407B9C71E4B39 CRC64;

Query Match 28.7%; Score 239; DB 2; Length 151;
Best Local Similarity 35.9%; Pred. No. 7.3e-14;
Matches 55; Conservative 28; Mismatches 54; Indels 16; Gaps 6;
QY 8 SLLIIISV---FLVGC--AQNFSRQEVGAATGAVVGVAGQFGKSGRVSMAIG-GAVL 61
Db 5 SKMIIIALAASMLQACNGPGGNKQGTGTLGGAGGALLGSGQFGKGGQL-VGVGVGALL 63
QY 62 GGLIGSKIGQSMDOQDK----IKLNOSLEKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYN 117
Db 64 GAVLGGQIVAGMDEQDRRLAELTSQRALEAAPSGSNVWRPNPDNGNYGYVTPNKTY---- 119
QY 118 QKERRQYCRFQOKAMTAGOKQBIYGTACPPQ 150
Db 120 -RNSTGQYCREYTTQTVVIGGKQKAYGNACRQ 151

RESULT 12

Q3F9Q1 PRELIMINARY; PRT; 131 AA.
ID Q3F9Q1;
AC Q3F9Q1;
DT 01-MAR-2001 (T-EMBLrel. 16, Created)
DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (T-EMBLrel. 22, Last annotation update)
DE 17 kDa protein (Fragment).
OS Rickettsia sp. California 2.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=147259;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=California 2;
RA Raoult D.;

RT "A new SFG rickettsia isolated from fleas.";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RC STRAIN=California 2;
 RA Roux V., Raoult D.;
 DR Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AF210693; AAG48554.1; -;
 FT NON_TER 1 1
 FT NON_TER 131 131
 SQ SEQUENCE 131 AA; 13374 MW; 23C8819B29FF860 CRC64;

Query Match 28.4%; Score 236.5; DB 2; Length 131;
 Best Local Similarity 38.3%; Pred. No. 1e-13;
 Matches 49; Conservative 24; Mismatches 44; Indels 11; Gaps 4;
 QY 25 SRQEVGAATGAVVGVAGQLFGKSGRVSMAIG-GAVLGGLIGSKIGSMDOODK----I 79
 Db 10 NKQGTGTLGGAGGALLGSQFGKGGQL-VGVGVGALLGAVLGQIGAGMDQDRRLAEL 68
 QY 80 KLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQERRQYCRFQOKAMIAGOK 139
 Db 69 TSQRALEATPSTSVENRPNPDNGNHYVTPNKTY-----RNSTGQYCREYTTQTVVIGGKQ 123
 QY 140 QBIYGTAC 147
 Db 124 QKAYGNAC 131

Query Match 28.4%; Score 236.5; DB 2; Length 131;
 Best Local Similarity 38.3%; Pred. No. 1e-13;
 Matches 49; Conservative 24; Mismatches 44; Indels 11; Gaps 4;
 QY 25 SRQEVGAATGAVVGVAGQLFGKSGRVSMAIG-GAVLGGLIGSKIGSMDOODK----I 79
 Db 10 NKQGTGTLGGAGGALLGSQFGKGGQL-VGVGVGALLGAVLGQIGAGMDQDRRLAEL 68
 QY 80 KLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQERRQYCRFQOKAMIAGOK 139
 Db 69 TSQRALEATPSTSVENRPNPDNGNHYVTPNKTY-----RNSTGQYCREYTTQTVVIGGKQ 123
 QY 140 QBIYGTAC 147
 Db 124 QKAYGNAC 131

Query Match 28.4%; Score 236.5; DB 2; Length 131;
 Best Local Similarity 38.3%; Pred. No. 1e-13;
 Matches 49; Conservative 24; Mismatches 44; Indels 11; Gaps 4;
 QY 25 SRQEVGAATGAVVGVAGQLFGKSGRVSMAIG-GAVLGGLIGSKIGSMDOODK----I 79
 Db 10 NKQGTGTLGGAGGALLGSQFGKGGQL-VGVGVGALLGAVLGQIGAGMDQDRRLAEL 68
 QY 80 KLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQERRQYCRFQOKAMIAGOK 139
 Db 69 TSQRALEATPSTSVENRPNPDNGNHYVTPNKTY-----RNSTGQYCREYTTQTVVIGGKQ 123
 QY 140 QBIYGTAC 147
 Db 124 QKAYGNAC 131

RESULT 13
 Q52637 ID Q52637 PRELIMINARY; PRT; 131 AA.
 AC Q52637;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DE 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
 DE 17 kDa antigen (Fragment).
 OS Rickettsia sp.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsiae; Rickettsia.
 OX NCBI_TaxID=789;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94117373; PubMed=8288533;
 RA Werren J.H., Hurst G.D., Zhang W., Breeuwer J.A., Stouthamer R.,
 RA Majerus M.E.;
 RT "Rickettsial relative associated with male killing in the ladybird
 beetle (*Adalia bipunctata*).";
 RL J. Bacteriol. 176:388-394 (1994).
 DR EMBL; U04162; AAA19235.1; -;
 FT NON_TER 1 1
 FT NON_TER 131 131
 SQ SEQUENCE 131 AA; 13344 MW; AIDCF71050DF52DF CRC64;

Query Match 28.4%; Score 236.5; DB 2; Length 131;
 Best Local Similarity 38.3%; Pred. No. 1e-13;
 Matches 49; Conservative 24; Mismatches 44; Indels 11; Gaps 4;
 QY 25 SRQEVGAATGAVVGVAGQLFGKSGRVSMAIG-GAVLGGLIGSKIGSMDOODK----I 79
 Db 10 NKQGTGTLGGAGGALLGSQFGKGGQL-VGVGVGALLGAVLGQIGAGMDQDRRLAEL 68
 QY 80 KLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQERRQYCRFQOKAMIAGOK 139
 Db 69 TSQRALEATPSTSVENRPNPDNGNHYVTPNKTY-----RNSTGQYCREYTTQTVVIGGKQ 123
 QY 140 QBIYGTAC 147
 Db 124 QKAYGNAC 131

Query Match 28.4%; Score 236.5; DB 2; Length 131;
 Best Local Similarity 38.3%; Pred. No. 1e-13;
 Matches 49; Conservative 24; Mismatches 44; Indels 11; Gaps 4;
 QY 25 SRQEVGAATGAVVGVAGQLFGKSGRVSMAIG-GAVLGGLIGSKIGSMDOODK----I 79
 Db 10 NKQGTGTLGGAGGALLGSQFGKGGQL-VGVGVGALLGAVLGQIGAGMDQDRRLAEL 68
 QY 80 KLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQERRQYCRFQOKAMIAGOK 139
 Db 69 TSQRALEATPSTSVENRPNPDNGNHYVTPNKTY-----RNSTGQYCREYTTQTVVIGGKQ 123
 QY 140 QBIYGTAC 147
 Db 124 QKAYGNAC 131

Query Match 28.4%; Score 236.5; DB 2; Length 131;
 Best Local Similarity 38.3%; Pred. No. 1e-13;
 Matches 49; Conservative 24; Mismatches 44; Indels 11; Gaps 4;
 QY 25 SRQEVGAATGAVVGVAGQLFGKSGRVSMAIG-GAVLGGLIGSKIGSMDOODK----I 79
 Db 10 NKQGTGTLGGAGGALLGSQFGKGGQL-VGVGVGALLGAVLGQIGAGMDQDRRLAEL 68
 QY 80 KLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQERRQYCRFQOKAMIAGOK 139
 Db 69 TSQRALEATPSTSVENRPNPDNGNHYVTPNKTY-----RNSTGQYCREYTTQTVVIGGKQ 123
 QY 140 QBIYGTAC 147
 Db 124 QKAYGNAC 131

Query Match 28.4%; Score 236.5; DB 2; Length 131;
 Best Local Similarity 38.3%; Pred. No. 1e-13;
 Matches 49; Conservative 24; Mismatches 44; Indels 11; Gaps 4;
 QY 25 SRQEVGAATGAVVGVAGQLFGKSGRVSMAIG-GAVLGGLIGSKIGSMDOODK----I 79
 Db 10 NKQGTGTLGGAGGALLGSQFGKGGQL-VGVGVGALLGAVLGQIGAGMDQDRRLAEL 68
 QY 80 KLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQERRQYCRFQOKAMIAGOK 139
 Db 69 TSQRALEATPSTSVENRPNPDNGNHYVTPNKTY-----RNSTGQYCREYTTQTVVIGGKQ 123
 QY 140 QBIYGTAC 147
 Db 124 QKAYGNAC 131

ID Q8KLU2 PRELIMINARY; PRT; 131 AA.
 AC Q8KLU2;
 DT 01-OCT-2002 (TRENBLrel. 22, Created)
 DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Outer membrane protein (Fragment).
 GN 17 kDa.
 OS Rickettsia helvetica.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsiae; Rickettsia.
 OX NCBI_TaxID=35789;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IrTAL;
 RX MEDLINE=22182650; PubMed=12194779;
 RA Beninati T., Lo N.L., Noda H., Eposito F., Rizzoli A., Favio G.,
 RA Genchi C.;
 RT "First detection of spotted fever group rickettsiae in Ixodes ricinus
 from Italy.";
 RL Emerg. Infect. Dis. 8:983-986 (2002).
 DR EMBL; AJ427881; CAD20877.1; -;
 FT NON_TER 1 1
 FT NON_TER 131 131
 SQ SEQUENCE 131 AA; 13383 MW; D78C171050CAA9CA CRC64;

Query Match 28.3%; Score 235.5; DB 2; Length 131;
 Best Local Similarity 38.3%; Pred. No. 1.3e-13;
 Matches 49; Conservative 24; Mismatches 44; Indels 11; Gaps 4;
 QY 25 SRQEVGAATGAVVGVAGQLFGKSGRVSMAIG-GAVLGGLIGSKIGSMDOODK----I 79
 Db 10 NKQGTGTLGGAGGALLGSQFGKGGQL-VGVGVGALLGAVLGQIGAGMDQDRRLAEL 68
 QY 80 KLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQERRQYCRFQOKAMIAGOK 139
 Db 69 TSQRALEATPSTSVENRPNPDNGNHYVTPNKTY-----RNSTGQYCREYTTQTVVIGGKQ 123
 QY 140 QBIYGTAC 147
 Db 124 QKAYGNAC 131

Query Match 28.3%; Score 235.5; DB 2; Length 131;
 Best Local Similarity 38.3%; Pred. No. 1.3e-13;
 Matches 49; Conservative 24; Mismatches 44; Indels 11; Gaps 4;
 QY 25 SRQEVGAATGAVVGVAGQLFGKSGRVSMAIG-GAVLGGLIGSKIGSMDOODK----I 79
 Db 10 NKQGTGTLGGAGGALLGSQFGKGGQL-VGVGVGALLGAVLGQIGAGMDQDRRLAEL 68
 QY 80 KLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQERRQYCRFQOKAMIAGOK 139
 Db 69 TSQRALEATPSTSVENRPNPDNGNHYVTPNKTY-----RNSTGQYCREYTTQTVVIGGKQ 123
 QY 140 QBIYGTAC 147
 Db 124 QKAYGNAC 131

RESULT 15
 Q9L522 ID Q9L522 PRELIMINARY; PRT; 131 AA.
 AC Q9L522;
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DE 17 kDa surface antigen (Fragment).
 OS Rickettsia peacockii.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsiae; Rickettsia.
 OX NCBI_TaxID=47589;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Dae100R;
 RX MEDLINE=21091941; PubMed=11157215;
 RA Simser J.A., Palmer A.T., Munderloh U.G., Kurti T.J.;
 RT "Isolation of a spotted fever group rickettsia, Rickettsia peacockii,
 in a Rocky Mountain wood tick, Dermacentor andersoni, cell line.";
 RL Appl. Environ. Microbiol. 67:546-552 (2001).
 DR EMBL; AF260571; AAF69012.1; -;
 FT NON_TER 1 1
 FT NON_TER 131 131
 SQ SEQUENCE 131 AA; 13413 MW; 228C020550CAA9D0 CRC64;

Query Match 28.3%; Score 235.5; DB 2; Length 131;
 Best Local Similarity 38.3%; Pred. No. 1.3e-13;
 Matches 49; Conservative 24; Mismatches 44; Indels 11; Gaps 4;
 QY 25 SRQEVGAATGAVVGVAGQLFGKSGRVSMAIG-GAVLGGLIGSKIGSMDOODK----I 79
 Db 10 NKQGTGTLGGAGGALLGSQFGKGGQL-VGVGVGALLGAVLGQIGAGMDQDRRLAEL 68
 QY 80 KLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQERRQYCRFQOKAMIAGOK 139
 Db 69 TSQRALEATPSTSVENRPNPDNGNHYVTPNKTY-----RNSTGQYCREYTTQTVVIGGKQ 123
 QY 140 QBIYGTAC 147
 Db 124 QKAYGNAC 131

Query Match 28.3%; Score 235.5; DB 2; Length 131;
 Best Local Similarity 38.3%; Pred. No. 1.3e-13;
 Matches 49; Conservative 24; Mismatches 44; Indels 11; Gaps 4;
 QY 25 SRQEVGAATGAVVGVAGQLFGKSGRVSMAIG-GAVLGGLIGSKIGSMDOODK----I 79
 Db 10 NKQGTGTLGGAGGALLGSQFGKGGQL-VGVGVGALLGAVLGQIGAGMDQDRRLAEL 68
 QY 80 KLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQERRQYCRFQOKAMIAGOK 139
 Db 69 TSQRALEATPSTSVENRPNPDNGNHYVTPNKTY-----RNSTGQYCREYTTQTVVIGGKQ 123
 QY 140 QBIYGTAC 147
 Db 124 QKAYGNAC 131

Query Match 28.3%; Score 235.5; DB 2; Length 131;
 Best Local Similarity 38.3%; Pred. No. 1.3e-13;
 Matches 49; Conservative 24; Mismatches 44; Indels 11; Gaps 4;
 QY 25 SRQEVGAATGAVVGVAGQLFGKSGRVSMAIG-GAVLGGLIGSKIGSMDOODK----I 79
 Db 10 NKQGTGTLGGAGGALLGSQFGKGGQL-VGVGVGALLGAVLGQIGAGMDQDRRLAEL 68
 QY 80 KLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQERRQYCRFQOKAMIAGOK 139
 Db 69 TSQRALEATPSTSVENRPNPDNGNHYVTPNKTY-----RNSTGQYCREYTTQTVVIGGKQ 123
 QY 140 QBIYGTAC 147
 Db 124 QKAYGNAC 131

Db 10 NKQGTGTLGGAGGALLGSQFGKGGQL-VGVGVGALLGAVLGQIGAGMDEQDRRLAEL 68
Qy 80 KLNQSLKVKAGQVTRWRNPDNGNSYSVEPVRTYQRYNKOERRQOYCREFQOKAMIAGQK 139
Db . 69 TSQEALETAPSGSNVENRPNNGNYGYTPNKTY-----RNSTGQYCREYTQTWVIGGKQ 123
Qy 140 QEIIYGTAC 147
Db 124 QKAYGNAC 131

Search completed: November 5, 2003, 20:15:04
Job time : 24.6356 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model
Run on: November 5, 2003, 20:12:57 ; Search time 137.938 Seconds
(without alignments)
1068.647 Million cell updates/sec

Title: US-09-677-374-2
Perfect score: 836
Sequence: 1 MNRGLOGSSLIISVFLVG.....IYGTACROPDGRWQVISTEK 162

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 5728757 seqs, 909918778 residues

Total number of hits satisfying chosen parameters: 5728757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents AA Main:*

1: /cgn2_6/ptodata/1/paa/ECTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US080_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097A_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US097B_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US099A_COMB.pep.*
25: /cgn2_6/ptodata/1/paa/US099B_COMB.pep.*
26: /cgn2_6/ptodata/1/paa/US100_COMB.pep.*
27: /cgn2_6/ptodata/1/paa/US101_COMB.pep.*
28: /cgn2_6/ptodata/1/paa/US102_COMB.pep.*
29: /cgn2_6/ptodata/1/paa/US103_COMB.pep.*
30: /cgn2_6/ptodata/1/paa/US104_COMB.pep.*
31: /cgn2_6/ptodata/1/paa/US106_COMB.pep.*
32: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	ID	Description
1	836	100.0	162 20	US-09-677-374-2 Sequence 2, Appli

2	836	100.0	162	28	US-10-261-446-2	Sequence 2, Appli
3	815	97.5	161	20	US-09-677-374-4	Sequence 4, Appli
4	815	97.5	161	28	US-10-261-446-4	Sequence 4, Appli
5	815	97.5	256	20	US-09-677-374-6	Sequence 6, Appli
6	815	97.5	256	28	US-10-261-446-6	Sequence 6, Appli
7	128.5	15.4	182	1	PCT-US02-12532-169	Sequence 169, App
8	128.5	15.4	182	27	US-10-127-032-169	Sequence 169, App
9	128.5	15.4	192	29	US-10-366-683-22817	Sequence 22817, A
10	128.5	15.4	192	30	US-10-419-128-22817	Sequence 22817, A
11	116	13.9	165	29	US-10-366-683-19701	Sequence 19701, A
12	116	13.9	165	30	US-10-419-128-19701	Sequence 19701, A
13	114.5	13.7	166	19	US-09-543-681A-4769	Sequence 4769, Ap
14	114.5	13.7	166	31	US-10-603-114-4769	Sequence 4769, Ap
15	113.5	13.6	155	22	US-09-791-537-85693	Sequence 85693, A
16	112	13.4	20	20	US-09-677-374-16	Sequence 16, Appl
17	112	13.4	20	28	US-10-261-446-16	Sequence 16, Appl
18	108	12.9	224	26	US-10-018-561-2	Sequence 2, Appli
19	108	12.9	224	26	US-10-018-561-4	Sequence 4, Appli
20	108	12.9	224	30	US-10-467-421-45	Sequence 45, Appl
21	108	12.9	231	19	US-09-540-236-3827	Sequence 3827, Ap
22	108	12.9	231	31	US-10-603-108-3827	Sequence 3827, Ap
23	108	12.9	231	32	US-60-128-476-3363	Sequence 3363, Ap
24	103	12.3	105	18	US-09-489-039A-10773	Sequence 10773, A
25	103	12.3	105	30	US-10-446-203-10773	Sequence 10773, A
26	102	12.2	154	22	US-09-791-537-117153	Sequence 117153,
27	102	12.2	155	28	US-10-274-586-319	Sequence 319, App
28	102	12.2	155	30	US-10-467-421-27	Sequence 27, Appl
29	101.5	12.1	306	29	US-10-366-683-21754	Sequence 21754, A
30	101.5	12.1	306	30	US-10-419-128-21754	Sequence 21754, A
31	100.5	12.0	83	21	US-09-739-449-12512	Sequence 12512, A
32	100.5	12.0	83	23	US-09-803-110-12512	Sequence 12512, A
33	100	12.0	199	18	US-09-489-039A-7592	Sequence 7592, Ap
34	100	12.0	199	30	US-10-446-203-7592	Sequence 7592, Ap
35	99	11.8	155	18	US-09-489-039A-9682	Sequence 9682, Ap
36	99	11.8	155	30	US-10-446-203-9682	Sequence 9682, Ap
37	98.5	11.8	387	29	US-10-366-683-23334	Sequence 23334, A
38	98.5	11.8	387	30	US-10-419-128-23334	Sequence 23334, A
39	98	11.7	175	22	US-09-758-472-9509	Sequence 9509, Ap
40	98	11.7	175	28	US-10-235-926-9509	Sequence 9509, Ap
41	98	11.7	203	16	US-09-252-691-7664	Sequence 7664, Ap
42	98	11.7	203	16	US-09-252-691C-7664	Sequence 7664, Ap
43	98	11.7	203	30	US-10-417-886-7664	Sequence 7664, Ap
44	98	11.7	309	1	PCT-US01-08631-46265	Sequence 46265, A
45	97.5	11.7	106	23	US-09-897-516-5127	Sequence 5127, Ap

ALIGNMENTS

RESULT 1
US-09-677-374-2
; Sequence 2, Application US/09677374
; GENERAL INFORMATION:
; APPLICANT: Burian, Jan
; APPLICANT: Kuzyk, Michael
; APPLICANT: Thornton, Julian
; APPLICANT: Kay, William
; TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY AGAINST
; FILE REFERENCE: RICKETTSIAL DISEASES, AND ASSOCIATED PREVENTATIVE THERAPY
; CURRENT APPLICATION NUMBER: US/09/677,374
; CURRENT FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/154,437
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: NO 20004637
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: IE 2000/0752
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: GB 0022825.4
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: CL 2544-2000
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 20

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Piscirickettsia salmonis
US-09-677-374-2
Query Match
Best Local Similarity 100.0%; Score 836; DB 20; Length 162;
Matches 162; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MNRGCLGSSLLIIISVFLVGCANFSEQVEVGAATGAVGVGAGQLFGKSGRVAIGA 60
Db 1 MNRGCLGSSLLIIISVFLVGCANFSEQVEVGAATGAVGVGAGQLFGKSGRVAIGA 60
QY 61 VLGLGSKIGSQMDQDDKIKLQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQ 120
Db 61 VLGLGSKIGSQMDQDDKIKLQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQ 120
QY 121 ERQOYCREFOQKAMIAQKQEIYGTACROPDGRWQVISTEK 162
Db 121 ERQOYCREFOQKAMIAQKQEIYGTACROPDGRWQVISTEK 162
RESULT 2
US-10-261-446-2
; Sequence 2, Application US/10261446
; GENERAL INFORMATION:
; APPLICANT: Kuzyk, Michael A.
; APPLICANT: Burian, Jan
; APPLICANT: Kay, William W.
; APPLICANT: Thornton, Julian C.
; TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY IN FISH AGAINST
; TITLE OF INVENTION: RICKETTSIAL DISEASES, AND ASSOCIATED PREVENTATIVE THERAPY
; FILE REFERENCE: 4616-64094
; CURRENT APPLICATION NUMBER: US/10/261,446
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US 09/677,374
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/154,437
; PRIOR FILING DATE: 1999-09-17
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Piscirickettsia salmonis
US-10-261-446-2
Query Match
Best Local Similarity 100.0%; Score 836; DB 28; Length 162;
Matches 162; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MNRGCLGSSLLIIISVFLVGCANFSEQVEVGAATGAVGVGAGQLFGKSGRVAIGA 60
Db 1 MNRGCLGSSLLIIISVFLVGCANFSEQVEVGAATGAVGVGAGQLFGKSGRVAIGA 60
QY 61 VLGLGSKIGSQMDQDDKIKLQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQ 120
Db 61 VLGLGSKIGSQMDQDDKIKLQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQ 120
QY 121 ERQOYCREFOQKAMIAQKQEIYGTACROPDGRWQVISTEK 162
Db 121 ERQOYCREFOQKAMIAQKQEIYGTACROPDGRWQVISTEK 162
RESULT 3
US-09-677-374-4
; Sequence 4, Application US/09677374
; GENERAL INFORMATION:
; APPLICANT: Burian, Jan
; APPLICANT: Kuzyk, Michael
; APPLICANT: Thornton, Julian
; TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY IN FISH AGAINST
; TITLE OF INVENTION: RICKETTSIAL DISEASES, AND ASSOCIATED PREVENTATIVE THERAPY
; FILE REFERENCE: 4616-64094
; CURRENT APPLICATION NUMBER: US/10/261,446
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US 09/677,374
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/154,437
; PRIOR FILING DATE: 1999-09-17
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Piscirickettsia salmonis
US-10-261-446-4
Query Match
Best Local Similarity 97.5%; Score 815; DB 20; Length 161;
Matches 158; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 3 RGCLOGSSLLIIISVFLVGCANFSEQVEVGAATGAVGVGAGQLFGKSGRVAIGA 62
Db 2 RGCLOGSSLLIIISVFLVGCANFSEQVEVGAATGAVGVGAGQLFGKSGRVAIGA 61
QY 63 GGLGSKIGSQMDQDDKIKLQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQ 122
Db 62 GGLGSKIGSQMDQDDKIKLQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQ 121
QY 123 ROQYCREFOQKAMIAQKQEIYGTACROPDGRWQVISTEK 162
Db 122 ROQYCREFOQKAMIAQKQEIYGTACROPDGRWQVISTEK 161
RESULT 4
US-10-261-446-4
; Sequence 4, Application US/10261446
; GENERAL INFORMATION:
; APPLICANT: Kuzyk, Michael A.
; APPLICANT: Burian, Jan
; APPLICANT: Kay, William W.
; APPLICANT: Thornton, Julian C.
; TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY IN FISH AGAINST
; TITLE OF INVENTION: RICKETTSIAL DISEASES, AND ASSOCIATED PREVENTATIVE THERAPY
; FILE REFERENCE: 4616-64094
; CURRENT APPLICATION NUMBER: US/10/261,446
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US 09/677,374
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/154,437
; PRIOR FILING DATE: 1999-09-17
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Piscirickettsia salmonis
US-10-261-446-4
Query Match
Best Local Similarity 97.5%; Score 815; DB 28; Length 161;
Matches 158; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 3 RGCLOGSSLLIIISVFLVGCANFSEQVEVGAATGAVGVGAGQLFGKSGRVAIGA 62
Db 2 RGCLOGSSLLIIISVFLVGCANFSEQVEVGAATGAVGVGAGQLFGKSGRVAIGA 61
QY 63 GGLGSKIGSQMDQDDKIKLQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQ 122
Db 62 GGLGSKIGSQMDQDDKIKLQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQ 121
QY 123 ROQYCREFOQKAMIAQKQEIYGTACROPDGRWQVISTEK 162
Db 122 ROQYCREFOQKAMIAQKQEIYGTACROPDGRWQVISTEK 161
```



```

; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US 09/677,374
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/154,437
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Piscirickettsia salmonis
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (-95)..(-1)
US-10-261-446-6

Query Match          97.5%; Score 815; DB 28; Length 256;
Best Local Similarity 98.8%; Pred. No. 1.5e-80;
Matches 158; Conservative 1; Mismatches 1; Indels 0; Gaps 0

QY      3  RGCLOGSSLLIIISVFLVGCQNFQRQEVGAATGAVVGGVAGQLFGKSGRVSMAIGGAVL 62
DB      3  RGCLOGSSLLIIISVFLVGCQNFQRQEVGAATGAVVGGVAGQLFGKSGRVSMAIGGAVL 156

QY      63  GGLIGSKIGQSMDDQDKIKLNQSLLEKVKKAGQVTRWRNPDGTGNSYSVEPVRTYQRYNKKQR 122
DB      157  GGLIGSKIGQSMDDQDKIKLNQSLLEKVKKAGQVTRWRNPDGTGNSYSVEPVRTYQRYNKKQR 216

QY      123  RQCYCREFOOKAMTAGOQOEIYGTACRQPDGRWQVISTEK 162
DB      217  RQCYCREFOOKAMTAGOQOEIYGTACRQPDGRWQVISTEK 256

RESULT 7
PCT-US02-12532-169
; Sequence 169, Application PC/TUS0212532
; GENERAL INFORMATION:
; APPLICANT: University of Iowa Research Foundation, Inc., et al.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF
; TITLE OF INVENTION: BIOFILM FORMATION
; FILE REFERENCE: UIZ-070CPPC
; CURRENT APPLICATION NUMBER: PCT/US02/12532
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/285,190
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: US 60/344,142
; PRIOR FILING DATE: 2001-10-24
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 169
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
PCT-US02-12532-169

Query Match          15.4%; Score 128.5; DB 1; Length 182;
Best Local Similarity 34.8%; Pred. No. 3.7e-05;
Matches 32; Conservative 13; Mismatches 28; Indels 19; Gaps 2

QY      28  QEVCAATGAVVGGVAGQLFGKSGRVSMAIGGAVLGLIGSKIGQSMDDQDKI----- 80
DB      70  QIAGTAGAVVGGLLGNQIGGTGKKTATVAGVGGVAGNKKVQSGMQERDYYTTTETRC 129

QY      81  -KLNQSLLEKVK-----KAGQVTRWRNP 100
DB      130  STVHDSSEKVGVDVQKYMMLDCKAGQIRMERDP 161

```

```
; APPLICANT: Bangera, M. Gita
; APPLICANT: Lory, Stephen
; APPLICANT: Greenberg, Everett Peter
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF
; FILE REFERENCE: BIOFILM FORMATION
; FILE REFERENCE: UIZ-070CP
; CURRENT APPLICATION NUMBER: US/10/127,032
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/285,190
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: US 60/344,142
; PRIOR FILING DATE: 2001-10-24
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 169
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-127-032-169

Query Match      15.4%; Score 128.5; DB 27; Length 182;
Best Local Similarity 34.8%; Pred. No. 3.7e-05;
Matches 32; Conservative 13; Mismatches 28; Indels 19; Gaps 2;

QY 28 QEVGAATGAVGVGAGQLFGKSGRVAMAIGGAVLGLIGSKIGQSMDOODKI----- 80
Db 70 QIAGTAIGAVVGGLLGNQIGGTGKKIATVAGAVGGYAGNKVQEGMQERDYYTTTETRC 129

QY 81 -KLNQSLKV-----KAGQVTRWRNP 100
Db 130 STVHDSSEKVGVDVKYMLDGGKAGQIRMERDP 161

RESULT 9
US-10-366-683-22817
; Sequence 22817, Application US/10366683
; GENERAL INFORMATION:
; APPLICANT: Rubenfield, Marc J.
; APPLICANT: Nolleng, Jork
; APPLICANT: Deloughery, Craig
; APPLICANT: Bush, David
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PATH03-04
; CURRENT APPLICATION NUMBER: US/10/366,683
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: 09/252,991
; PRIOR FILING DATE: 1999-02-18
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22817
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-366-683-22817

Query Match      15.4%; Score 128.5; DB 29; Length 192;
Best Local Similarity 34.8%; Pred. No. 4e-05;
Matches 32; Conservative 13; Mismatches 28; Indels 19; Gaps 2;

QY 28 QEVGAATGAVGVGAGQLFGKSGRVAMAIGGAVLGLIGSKIGQSMDOODKI----- 80
Db 80 QIAGTAIGAVVGGLLGNQIGGTGKKIATVAGAVGGYAGNKVQEGMQERDYYTTTETRC 139

QY 81 -KLNQSLKV-----KAGQVTRWRNP 100
Db 140 STVHDSSEKVGVDVKYMLDGGKAGQIRMERDP 171

RESULT 10
US-10-419-128-22817
; Sequence 22817, Application US/10419128
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
```

```
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/10/419,128
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: US/09/252,991
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22817
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-419-128-22817

Query Match      15.4%; Score 128.5; DB 30; Length 192;
Best Local Similarity 34.8%; Pred. No. 4e-05;
Matches 32; Conservative 13; Mismatches 28; Indels 19; Gaps 2;

QY 28 QEVGAATGAVGVGAGQLFGKSGRVAMAIGGAVLGLIGSKIGQSMDOODKI----- 80
Db 80 QIAGTAIGAVVGGLLGNQIGGTGKKIATVAGAVGGYAGNKVQEGMQERDYYTTTETRC 139

QY 81 -KLNQSLKV-----KAGQVTRWRNP 100
Db 140 STVHDSSEKVGVDVKYMLDGGKAGQIRMERDP 171

RESULT 11
US-10-366-683-19701
; Sequence 19701, Application US/10366683
; GENERAL INFORMATION:
; APPLICANT: Rubenfield, Marc J.
; APPLICANT: Nolleng, Jork
; APPLICANT: Deloughery, Craig
; APPLICANT: Bush, David
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PATH03-04
; CURRENT APPLICATION NUMBER: US/10/366,683
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: 09/252,991
; PRIOR FILING DATE: 1999-02-18
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19701
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-366-683-19701

Query Match      13.9%; Score 116; DB 29; Length 165;
Best Local Similarity 25.0%; Pred. No. 0.00078;
Matches 34; Conservative 21; Mismatches 41; Indels 40; Gaps 3;

QY 1 MNRGCGQSSLLIIISVFLVGC-----AQNFSPQE----- 29
Db 12 MKRSALIVASFAMALAGGCCOSSITGDTYSREEARVTVTVMGTIOALRPVKIEGKTP 71
QY 30 VGAATGAVVGGVAGQLFGKSGRVAMAIGGAVLGLIGSKIGQSMDOODKI----- 82
Db 72 IGSIAAGVGGVAGSAVGGGKGYAAIIGAAGLGAATEGLTRTQGVITVREDGG 131

QY 83 --NQSLKVAGQVTR 96
Db 132 STRAYVQVDQGIIFR 147

RESULT 12
US-10-419-128-19701
; Sequence 19701, Application US/10419128
```

GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: ABRUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/10/419,128
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: US/09/252,991
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19701
; LENGTH: 165
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-419-128-19701

Query Match 13.9%; Score 116; DB 30; Length 165;
Best Local Similarity 25.0%; Pred. No. 0.00078;
Matches 34; Conservative 21; Mismatches 41; Indels 40; Gaps 3;

QY 1 MNRGCGSSLLIIISVFLVGC-----AQNFSRQE----- 29
Db 12 MKKSALIVASFTAMALALGCGQSSLTGDTYSREEARVTQVRMTIQALRPVKIEGTKTP 71
QY 30 VCAATGAVVGGVAGQVFGKSGRVAIAGGAVLGGIGSKIGSQMDQDKIKL----- 82
Db 72 IGSIAAGVGGVAGSAGVGGKGSVAAIIGAVAGLLGAATEGLTRTQGVBITVREDDG 131
QY 83 --NQSLEKVKAGQVTR 96
Db 132 STRAVVQVDQGIIFR 147

RESULT 13
US-09-543-681A-4769
; Sequence 4769, Application US/09543681A
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4769
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4769

Query Match 13.7%; Score 114.5; DB 19; Length 166;
Best Local Similarity 25.4%; Pred. No. 0.0011;
Matches 35; Conservative 19; Mismatches 53; Indels 31; Gaps 3;

QY 4 GCGQSSSL-----IIISVFLVGCQAQNFQSRQEVGAATGAVVGGVAGQ 44
Db 28 GCVNTSSLSGDTYTSQAQQAQNVYTGTVSVRAVNIQAGSDENVLGAIGGAVLGGLLGN 87
QY 45 LFGKSGRVAIAGGAVLGGIGSKIGSQMDQDKIKL-----VKAGQVTRWR 98
Db 88 TTGGTGRNLATAAGAIAGMAGQAQAGALNTTKGVLEVRDLSGKTVVVGKADNTAYR 147
QY 99 NPD-----TGNSYSVEP 110
Db 148 QGQRVAVIGNGNLTVSP 165

US-09-543-681A-4769
; Sequence 4769, Application US/09543681A
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4769
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4769

Query Match 13.7%; Score 114.5; DB 19; Length 166;
Best Local Similarity 25.4%; Pred. No. 0.0011;
Matches 35; Conservative 19; Mismatches 53; Indels 31; Gaps 3;

QY 4 GCGQSSSL-----IIISVFLVGCQAQNFQSRQEVGAATGAVVGGVAGQ 44
Db 28 GCVNTSSLSGDTYTSQAQQAQNVYTGTVSVRAVNIQAGSDENVLGAIGGAVLGGLLGN 87
QY 45 LFGKSGRVAIAGGAVLGGIGSKIGSQMDQDKIKL-----VKAGQVTRWR 98
Db 88 TTGGTGRNLATAAGAIAGMAGQAQAGALNTTKGVLEVRDLSGKTVVVGKADNTAYR 147
QY 99 NPD-----TGNSYSVEP 110
Db 148 QGQRVAVIGNGNLTVSP 165

RESULT 14
US-10-603-114-4769
; Sequence 4769, Application US/10603114
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/10/603,114
; CURRENT FILING DATE: 2003-06-24
; PRIOR APPLICATION NUMBER: US/09/543,681
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4769
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-10-603-114-4769

Query Match 13.7%; Score 114.5; DB 31; Length 166;
Best Local Similarity 25.4%; Pred. No. 0.0011;
Matches 35; Conservative 19; Mismatches 53; Indels 31; Gaps 3;

QY 4 GCGQSSSL-----IIISVFLVGCQAQNFQSRQEVGAATGAVVGGVAGQ 44
Db 28 GCVNTSSLSGDTYTSQAQQAQNVYTGTVSVRAVNIQAGSDENVLGAIGGAVLGGLLGN 87
QY 45 LFGKSGRVAIAGGAVLGGIGSKIGSQMDQDKIKL-----VKAGQVTRWR 98
Db 88 TTGGTGRNLATAAGAIAGMAGQAQAGALNTTKGVLEVRDLSGKTVVVGKADNTAYR 147
QY 99 NPD-----TGNSYSVEP 110
Db 148 QGQRVAVIGNGNLTVSP 165

RESULT 15
US-09-791-537-85693
; Sequence 85693, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 85693
; LENGTH: 155
; TYPE: PRT
; ORGANISM: Yersinia enterocolitica
US-09-791-537-85693

Query Match 13.6%; Score 113.5; DB 22; Length 155;
Best Local Similarity 24.7%; Pred. No. 0.0013;
Matches 37; Conservative 19; Mismatches 43; Indels 51; Gaps 4;

QY 10 SLIIISVFLVGCQAQNF-----FSRQE-----VGAAT 34
Db 7 AVAIAAATLTGCANNNTLSGDFVSASQAQVQTVTYTGLLSVPRVTIQGGDDNNVMAIG 66
QY 35 GAVVGGVAGQVFGKSGRVAIAGGAVLGGIGSKIGSQMDQDKIKL----- 82
Db 67 GAVLGGFLGNTVGGGTGRSLATAAGAVAGMAGQVQAGMNRDTGVQLEVRKDDGTTILV 126
QY 83 --NQSLEKVKAGQVTRWRNPDTGNSYSVEP 110
Db 127 VQKQGPTRFSVQ--RVMLASSGSTVTVSP 154

Search completed: November 5, 2003, 20:25:12
Job time : 138.938 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 20:13:32 ; Search time 9.23316 Seconds
(without alignments)
824.821 Million cell updates/sec

Title: US-09-677-374-2

Perfect score: 836

Sequence: 1 MNRGCLQSSLLIIISVFLVG.....IYGTACRQPDGRWQVISTEK 162

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 248976 segs, 47010500 residues

Total number of hits satisfying chosen parameters: 248976

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New.*

- 1: /cgn2_6/prodata/2/paa/PCT_NEW_COMB.pep.*
- 2: /cgn2_6/prodata/2/paa/US06_NEW_COMB.pep.*
- 3: /cgn2_6/prodata/2/paa/US07_NEW_COMB.pep.*
- 4: /cgn2_6/prodata/2/paa/US08_NEW_COMB.pep.*
- 5: /cgn2_6/prodata/2/paa/US09_NEW_COMB.pep.*
- 6: /cgn2_6/prodata/2/paa/US10_NEW_COMB.pep.*
- 7: /cgn2_6/prodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	836	100.0	162	US-10-241-602B-6	Sequence 6, Appli
2	836	100.0	162	US-10-261-445B-2	Sequence 2, Appli
3	815	97.5	161	US-10-261-445B-4	Sequence 4, Appli
4	815	97.5	256	US-10-261-445B-6	Sequence 6, Appli
5	112	13.4	20	US-10-261-445B-16	Sequence 16, Appl
6	105.5	12.6	223	US-09-581-286A-434	Sequence 434, App
7	105.5	12.6	230	US-09-581-286A-309	Sequence 309, App
8	97.5	11.7	106	US-09-897-516A-5133	Sequence 5133, App
9	83.5	10.0	99	US-10-389-647-605	Sequence 605, App
10	81.5	9.7	734	US-10-425-114A-49440	Sequence 49440, A
11	80	9.6	403	US-09-897-516A-7766	Sequence 7766, App
12	80	9.6	515	US-10-679-063-22745	Sequence 22745, A
13	80	9.6	517	US-10-679-063-22744	Sequence 22744, A
14	80	9.6	1755	1 PCT-US03-26780-3444	Sequence 3444, App
15	80	9.6	2599	US-09-674-546A-1669	Sequence 1669, App
16	79	9.4	593	US-10-435-696-50	Sequence 50, Appl
17	79	9.4	1261	US-10-679-063-15542	Sequence 15542, A
18	76	9.1	293	1 PCT-US03-28227-2863	Sequence 2863, App
19	76	9.1	423	1 PCT/US02/39429-8	Sequence 8, Appli
20	74	8.9	800	US-10-296-115-1229	Sequence 1229, App
21	74	8.9	820	1 PCT-US03-28227-5403	Sequence 5403, App
22	74	8.9	871	1 PCT-US03-31442-3	Sequence 3, Appli
23	73.5	8.8	390	1 PCT-US02-38594-98	Sequence 98, Appl
24	73.5	8.8	390	1 PCT-US02-38594-100	Sequence 100, App
25	73.5	8.8	404	1 PCT-US02-38594-96	Sequence 96, Appl
26	73.5	8.8	404	US-60-487-610-2773	Sequence 2773, App

27	73.5	8.8	733	5	US-09-743-818A-4	Sequence 4, Appli
28	73	8.7	143	6	US-10-425-114A-70833	Sequence 70833, A
29	73	8.7	254	1	PCT-US03-10856A-5	Sequence 5, Appli
30	73	8.7	2322	1	PCT-US03-26780-2711	Sequence 2711, App
31	72.5	8.7	261	6	US-10-689-006-24	Sequence 24, Appl
32	72.5	8.7	515	5	US-09-743-818A-71	Sequence 71, Appl
33	72.5	8.7	571	5	US-09-743-818A-7	Sequence 7, Appli
34	72.5	8.7	692	1	PCT-US02-24483-40	Sequence 40, Appl
35	72.5	8.7	698	5	US-09-743-818A-5	Sequence 5, Appli
36	72.5	8.7	711	1	PCT-US02-24483-38	Sequence 38, Appl
37	72.5	8.7	757	7	US-60-487-610-1796	Sequence 1796, App
38	72	8.6	89	6	US-10-472-928-932	Sequence 932, App
39	72	8.6	101	1	PCT-US03-32968-3	Sequence 3, Appli
40	72	8.6	101	1	PCT-US03-32968-6	Sequence 6, Appli
41	72	8.6	399	5	US-09-897-516A-4405	Sequence 4405, App
42	72	8.6	405	4	US-08-592-070D-14	Sequence 14, Appl
43	72	8.6	512	6	US-10-679-063-22287	Sequence 22287, A
44	72	8.6	541	7	US-60-487-610-2783	Sequence 2783, App
45	72	8.6	563	7	US-60-487-610-1525	Sequence 1525, App

ALIGNMENTS

RESULT 1
US-10-241-602B-6
; Sequence 6, Application US/10241602B
; GENERAL INFORMATION:
; APPLICANT: Simard, Nathalie
; APPLICANT: Brouwers, Huub
; APPLICANT: Jones, Simon
; APPLICANT: Griffiths, Steve
; APPLICANT: Valenzuela, Pablo
; APPLICANT: Burzio, Luis
; TITLE OF INVENTION: Sequences from Piscirickettsia salmonis
; FILE REFERENCE: H-32319A
; CURRENT APPLICATION NUMBER: US/10/241,602B
; PRIOR FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: PCT/GB01/01055
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: GB0005838.8
; PRIOR FILING DATE: 2000-03-11
; PRIOR APPLICATION NUMBER: GB0016080.4
; PRIOR FILING DATE: 2000-07-01
; PRIOR APPLICATION NUMBER: GB0016082.0
; PRIOR FILING DATE: 2000-07-01
; PRIOR APPLICATION NUMBER: GB0018599.1
; PRIOR FILING DATE: 2000-07-29
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Piscirickettsia salmonis
US-10-241-602B-6

Query Match	100.0%	Score 836;	DB 6;	Length 162;
Best Local Similarity	100.0%	Pred. No. 6.4e-65;	Mismatches 0;	Indels 0;
Matches 162;	Conservative	0;		Gaps 0;
QY	1	MNRGCLQSSLLIIISVFLVGCAQNFNRQEVGAGTAAVVGAGQJFGKSGRVAMAIGA	60	
Db	1	MNRGCLQSSLLIIISVFLVGCAQNFNRQEVGAGTAAVVGAGQJFGKSGRVAMAIGA	60	
QY	61	VLGGLIGSKIGSMDQDKIKLNQSLKVKVAGQVTRWRNPDTGNSYSVEPVTYQRYNQ	120	
Db	61	VLGGLIGSKIGSMDQDKIKLNQSLKVKVAGQVTRWRNPDTGNSYSVEPVTYQRYNQ	120	
QY	121	ERRQYCFRFOQKAMIAQOKBIYGTACRQPDGRWQVISTEK	162	
Db	121	ERRQYCFRFOQKAMIAQOKBIYGTACRQPDGRWQVISTEK	162	

```
RESULT 2
US-10-261-445B-2
; Sequence 2, Application US/10261445B
; GENERAL INFORMATION:
; APPLICANT: Kuzyk, Michael A.
; APPLICANT: Burian, Jan
; APPLICANT: Kay, William W.
; APPLICANT: Thornton, Julian C.
; TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY IN FISH AGAINST
; TITLE OF INVENTION: RICKETTSIAL DISEASES, AND ASSOCIATED PREVENTATIVE THERAPY
; FILE REFERENCE: 4616-64095
; CURRENT APPLICATION NUMBER: US/10/261,445B
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US 09/677,374
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/154,437
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Piscirickettsia salmonis
US-10-261-445B-2

Query Match      100.0%; Score 836; DB 6; Length 162;
Best Local Similarity 100.0%; Pred. No. 6.4e-65;
Matches 162; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGCLQGSSLLIIISVFLVGCQNFSRQEVGAATGAVVGGVAGQLFGKSGRVAIGA 60
Db 1 MNRGCLQGSSLLIIISVFLVGCQNFSRQEVGAATGAVVGGVAGQLFGKSGRVAIGA 60
QY 61 VLGLGSKIGQSMDOODKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQYNKQ 120
Db 61 VLGLGSKIGQSMDOODKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQYNKQ 120
QY 121 ERQQYCRFFQOKAMIAQKQEIYGTACRQPDGRWQVISTEK 162
Db 121 ERQQYCRFFQOKAMIAQKQEIYGTACRQPDGRWQVISTEK 162

RESULT 3
US-10-261-445B-4
; Sequence 4, Application US/10261445B
; GENERAL INFORMATION:
; APPLICANT: Kuzyk, Michael A.
; APPLICANT: Burian, Jan
; APPLICANT: Kay, William W.
; APPLICANT: Thornton, Julian C.
; TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY IN FISH AGAINST
; TITLE OF INVENTION: RICKETTSIAL DISEASES, AND ASSOCIATED PREVENTATIVE THERAPY
; FILE REFERENCE: 4616-64095
; CURRENT APPLICATION NUMBER: US/10/261,445B
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US 09/677,374
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/154,437
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Piscirickettsia salmonis
US-10-261-445B-4

Query Match      97.5%; Score 815; DB 6; Length 161;
Best Local Similarity 98.8%; Pred. No. 4.1e-63;
Matches 158; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 RGCLOGSSLLIIISVFLVGCQNFSRQEVGAATGAVVGGVAGQLFGKSGRVAIGA 62
Db 3 RGCLOGSSLLIIISVFLVGCQNFSRQEVGAATGAVVGGVAGQLFGKSGRVAIGA 62
QY 97 RGLGSKIGQSMDOODKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQYNKQ 122
Db 97 RGLGSKIGQSMDOODKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQYNKQ 122
QY 157 GGLGSKIGQSMDOODKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQYNKQ 216
Db 157 GGLGSKIGQSMDOODKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQYNKQ 216

RESULT 4
US-10-261-445B-6
; Sequence 6, Application US/10261445B
; GENERAL INFORMATION:
; APPLICANT: Kuzyk, Michael A.
; APPLICANT: Burian, Jan
; APPLICANT: Kay, William W.
; APPLICANT: Thornton, Julian C.
; TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY IN FISH AGAINST
; TITLE OF INVENTION: RICKETTSIAL DISEASES, AND ASSOCIATED PREVENTATIVE THERAPY
; FILE REFERENCE: 4616-64095
; CURRENT APPLICATION NUMBER: US/10/261,445B
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US 09/677,374
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/154,437
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Piscirickettsia salmonis
; NAME/KEY: SIGNAL
; LOCATION: (-95)..(-1)
US-10-261-445B-6

Query Match      97.5%; Score 815; DB 6; Length 256;
Best Local Similarity 98.8%; Pred. No. 7.1e-63;
Matches 158; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 RGCLOGSSLLIIISVFLVGCQNFSRQEVGAATGAVVGGVAGQLFGKSGRVAIGA 62
Db 3 RGCLOGSSLLIIISVFLVGCQNFSRQEVGAATGAVVGGVAGQLFGKSGRVAIGA 62
QY 97 RGLGSKIGQSMDOODKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQYNKQ 122
Db 97 RGLGSKIGQSMDOODKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQYNKQ 122
QY 157 GGLGSKIGQSMDOODKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQYNKQ 216
Db 157 GGLGSKIGQSMDOODKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQYNKQ 216

RESULT 5
US-10-261-445B-16
; Sequence 16, Application US/10261445B
; GENERAL INFORMATION:
; APPLICANT: Kuzyk, Michael A.
; APPLICANT: Burian, Jan
; APPLICANT: Kay, William W.
; APPLICANT: Thornton, Julian C.
; TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY IN FISH AGAINST
; TITLE OF INVENTION: RICKETTSIAL DISEASES, AND ASSOCIATED PREVENTATIVE THERAPY
; FILE REFERENCE: 4616-64095
; CURRENT APPLICATION NUMBER: US/10/261,445B
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US 09/677,374
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/154,437
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Piscirickettsia salmonis
US-10-261-445B-16

Query Match      97.5%; Score 815; DB 6; Length 256;
Best Local Similarity 98.8%; Pred. No. 7.1e-63;
Matches 158; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 RGCLOGSSLLIIISVFLVGCQNFSRQEVGAATGAVVGGVAGQLFGKSGRVAIGA 62
Db 3 RGCLOGSSLLIIISVFLVGCQNFSRQEVGAATGAVVGGVAGQLFGKSGRVAIGA 62
QY 97 RGLGSKIGQSMDOODKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQYNKQ 122
Db 97 RGLGSKIGQSMDOODKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQYNKQ 122
QY 157 GGLGSKIGQSMDOODKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQYNKQ 216
Db 157 GGLGSKIGQSMDOODKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQYNKQ 216
```

```
Db 2 RGCLOGSSLLIIISVFLVGCQNFSRQEVGAATGAVVGGVAGQLFGKSGRVAIGA 61
QY 63 GGLGSKIGQSMDOODKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQYNKQ 122
Db 62 GGLGSKIGQSMDOODKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQYNKQ 121
QY 123 RQCYCREPQOKAMIAQKQEIYGTACRQPDGRWQVISTEK 162
Db 122 RQCYCREPQOKAMIAQKQEIYGTACRQPDGRWQVISTEK 161

RESULT 4
US-10-261-445B-6
; Sequence 6, Application US/10261445B
; GENERAL INFORMATION:
; APPLICANT: Kuzyk, Michael A.
; APPLICANT: Burian, Jan
; APPLICANT: Kay, William W.
; APPLICANT: Thornton, Julian C.
; TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY IN FISH AGAINST
; TITLE OF INVENTION: RICKETTSIAL DISEASES, AND ASSOCIATED PREVENTATIVE THERAPY
; FILE REFERENCE: 4616-64095
; CURRENT APPLICATION NUMBER: US/10/261,445B
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US 09/677,374
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/154,437
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Piscirickettsia salmonis
; NAME/KEY: SIGNAL
; LOCATION: (-95)..(-1)
US-10-261-445B-6

Query Match      97.5%; Score 815; DB 6; Length 256;
Best Local Similarity 98.8%; Pred. No. 7.1e-63;
Matches 158; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 RGCLOGSSLLIIISVFLVGCQNFSRQEVGAATGAVVGGVAGQLFGKSGRVAIGA 62
Db 3 RGCLOGSSLLIIISVFLVGCQNFSRQEVGAATGAVVGGVAGQLFGKSGRVAIGA 62
QY 97 RGLGSKIGQSMDOODKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQYNKQ 122
Db 97 RGLGSKIGQSMDOODKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQYNKQ 122
QY 157 GGLGSKIGQSMDOODKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQYNKQ 216
Db 157 GGLGSKIGQSMDOODKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQYNKQ 216

RESULT 5
US-10-261-445B-16
; Sequence 16, Application US/10261445B
; GENERAL INFORMATION:
; APPLICANT: Kuzyk, Michael A.
; APPLICANT: Burian, Jan
; APPLICANT: Kay, William W.
; APPLICANT: Thornton, Julian C.
; TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY IN FISH AGAINST
; TITLE OF INVENTION: RICKETTSIAL DISEASES, AND ASSOCIATED PREVENTATIVE THERAPY
; FILE REFERENCE: 4616-64095
; CURRENT APPLICATION NUMBER: US/10/261,445B
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US 09/677,374
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/154,437
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Piscirickettsia salmonis
US-10-261-445B-16

Query Match      97.5%; Score 815; DB 6; Length 256;
Best Local Similarity 98.8%; Pred. No. 7.1e-63;
Matches 158; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 RGCLOGSSLLIIISVFLVGCQNFSRQEVGAATGAVVGGVAGQLFGKSGRVAIGA 62
Db 3 RGCLOGSSLLIIISVFLVGCQNFSRQEVGAATGAVVGGVAGQLFGKSGRVAIGA 62
QY 97 RGLGSKIGQSMDOODKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQYNKQ 122
Db 97 RGLGSKIGQSMDOODKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQYNKQ 122
QY 157 GGLGSKIGQSMDOODKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQYNKQ 216
Db 157 GGLGSKIGQSMDOODKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQYNKQ 216
```



```
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8415
; SEQ ID NO 5133
; LENGTH: 106
; TYPE: PRT
; ORGANISM: Xenorhabdus sp.
US-09-897-516A-5133

Query Match      11.7%; Score 97.5; DB 5; Length 106;
Best Local Similarity 36.4%; Pred. No. 0.12;
Matches 28; Conservative 8; Mismatches 32; Indels 9; Gaps 2;

QY 28 QEVGAATGAVGVGAGQLFKGS--GRVMAIGAVGLGLIGSKI-----GQSMDOOD 78
Db 4 QTFGAILGAVAGSVGVNAGSGGAGTAAGAVGGGALGAAGSVVKDKVIVEGVSUTYKE 63

QY 79 KIKLNOSLEKVKAGQVT 95
Db 64 KTKIYTSTQVGRTCQFT 80

RESULT 9
US-10-389-647-605
; Sequence 605, Application US/10389647
; GENERAL INFORMATION:
; APPLICANT: GREENBERG, E. Peter
; APPLICANT: SCHUSTER, Martin
; APPLICANT: LOSTROH, Candi
; TITLE OF INVENTION: QUORUM SENSING SIGNALING IN BACTERIA
; FILE REFERENCE: UIZ-038CP
; CURRENT APPLICATION NUMBER: US/10/389,647
; CURRENT FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: 09/653730
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/153022
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 710
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 605
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-389-647-605

Query Match      10.0%; Score 83.5; DB 6; Length 99;
Best Local Similarity 32.9%; Pred. No. 1.8;
Matches 23; Conservative 9; Mismatches 21; Indels 17; Gaps 3;

QY 5 CLOGSSLIITISVPLVCCANFRQEVGAATGAVGVGAGQLFKGSGRVMAIG----- 58
Db 11 CVSGGTL--SGMIVGAVD-----GAATGMAIGKGGAGGFGGALSQLVLIVPTA 60

QY 59 -GAVLGGLIG 67
Db 61 MGAAGGTGV 70

RESULT 10
US-10-425-114A-49440
; Sequence 49440, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 49440
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700991377_FLI.pep
US-10-425-114A-49440

Query Match      9.7%; Score 81.5; DB 6; Length 734;
Best Local Similarity 25.5%; Pred. No. 29;
Matches 35; Conservative 24; Mismatches 49; Indels 29; Gaps 6;

QY 26 SRQEVGAATGAVVGVA--GQLFGKGSGRV-----AMAIGGAVLG---GLIGSKI 70
Db 77 SKNEVNGVAFADGNGVTEHGEVWGSKBVAVNNEVAIVDNGVAYGGEVHGSKNGAVNNEV 136

QY 71 -----GQSMDOODKIKLNOSLEKVKAGQVTRWRNPDTCNSYSVEPVRTYQRYNKQERQ 124
Db 137 VIADGNGVTEGQEDHCLKNETVNNVVA-----NADEGNSGAVECFQTYKRRKHAKSSS 189

QY 125 QY-CREFOQKAMIAGQK 140
Db 190 EFKVQENSRRKHGAASQ 206

RESULT 11
US-09-897-516A-7766
; Sequence 7766, Application US/09897516A
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215,161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8415
; SEQ ID NO 7766
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Xenorhabdus sp.
US-09-897-516A-7766

Query Match      9.6%; Score 80; DB 5; Length 403;
Best Local Similarity 34.2%; Pred. No. 19;
Matches 26; Conservative 11; Mismatches 33; Indels 6; Gaps 3;

QY 31 GAATGAVGVGAGQLFK-KGSGRVMAIG--GAVLGLIGSKIGQSMDOODKIKLNQSL 87
Db 137 GHGTGAAGKVFTRYGFENATEVAMACATFGLILGLIGGFPVRLVRNDK---TPGLE 193

QY 88 KVKAGQVTRWRNPDGT 103
Db 194 NDDAEVPSAFEKPYTG 209

RESULT 12
US-10-679-063-22745
; Sequence 22745, Application US/10679063
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
```


FILE REFERENCE: 38-15(52054)B
 CURRENT APPLICATION NUMBER: US/10/679,063
 CURRENT FILING DATE: 2003-10-02
 PRIOR APPLICATION NUMBER: 60/415,758
 PRIOR FILING DATE: 2002-10-02
 NUMBER OF SEQ ID NOS: 27373
 SEQ ID NO 22745
 LENGTH: 515
 TYPE: PRT
 ORGANISM: Anopheles gambiae str. PEST
 US-10-679-063-22745

Query Match 9.6%; Score 80; DB 6; Length 515;
 Best Local Similarity 28.1%; Pred. No. 25;
 Matches 25; Conservative 15; Mismatches 41; Indels 8; Gaps 2;
 QY 41 VAGQLFGKSGRVAMA-----IGGAVLGGIGKIGQSMDDQDKIK-LNOSLEKVKAG 92
 DB 45 VAGNHEGGMKVAARRPVLGELGNKVLNRNASQDLGKGVKGAALKNANPTLKNIKPR 104
 QY 93 QVTRWRNPDTGNSYSVEPVRTYQRYNKQE 121
 DB 105 VDTWRKADTAAAVVPKVVTRSDSQK 133

RESULT 13
 US-10-679-063-22744
 Sequence 22744, Application US/10679063
 GENERAL INFORMATION:
 APPLICANT: Edgerton, Michael D
 TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
 FILE REFERENCE: 38-15(52054)B
 CURRENT APPLICATION NUMBER: US/10/679,063
 CURRENT FILING DATE: 2003-10-02
 PRIOR APPLICATION NUMBER: 60/415,758
 PRIOR FILING DATE: 2002-10-02
 NUMBER OF SEQ ID NOS: 27373
 SEQ ID NO 22744
 LENGTH: 517
 TYPE: PRT
 ORGANISM: Anopheles gambiae str. PEST
 US-10-679-063-22744

Query Match 9.6%; Score 80; DB 6; Length 517;
 Best Local Similarity 28.1%; Pred. No. 25;
 Matches 25; Conservative 15; Mismatches 41; Indels 8; Gaps 2;
 QY 41 VAGQLFGKSGRVAMA-----IGGAVLGGIGKIGQSMDDQDKIK-LNOSLEKVKAG 92
 DB 45 VAGNHEGGMKVAARRPVLGELGNKVLNRNASQDLGKGVKGAALKNANPTLKNIKPR 104
 QY 93 QVTRWRNPDTGNSYSVEPVRTYQRYNKQE 121
 DB 105 VDTWRKADTAAAVVPKVVTRSDSQK 133

RESULT 14
 PCT-US03-26780-3444
 Sequence 3444, Application PC/TUS0326780
 GENERAL INFORMATION:
 APPLICANT: FIVEPRIME THERAPEUTICS, INC.
 TITLE OF INVENTION: HUMAN POLYPEPTIDES ENCODED BY POLYNUCLEOTIDES AND METHODS OF
 FILE REFERENCE: 08940.0014-00304
 CURRENT APPLICATION NUMBER: PCT/US03/26780
 CURRENT FILING DATE: 2003-08-28
 PRIOR APPLICATION NUMBER: 60/406,616
 PRIOR FILING DATE: 2002-08-29
 PRIOR APPLICATION NUMBER: 60/406,579
 PRIOR FILING DATE: 2002-08-29
 PRIOR APPLICATION NUMBER: 60/406,655
 PRIOR FILING DATE: 2002-08-29
 PRIOR APPLICATION NUMBER: 60/406,642

PRIOR FILING DATE: 2002-08-29
 PRIOR APPLICATION NUMBER: 60/406,640
 PRIOR FILING DATE: 2002-08-29
 PRIOR APPLICATION NUMBER: 60/406,588
 PRIOR FILING DATE: 2002-08-29
 PRIOR APPLICATION NUMBER: 60/406,576
 PRIOR FILING DATE: 2002-08-29
 PRIOR APPLICATION NUMBER: 60/406,646
 PRIOR FILING DATE: 2002-08-29
 PRIOR APPLICATION NUMBER: 60/406,666
 PRIOR FILING DATE: 2002-08-29
 PRIOR APPLICATION NUMBER: 60/406,653
 PRIOR FILING DATE: 2002-08-29
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 3700
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 3444
 LENGTH: 1755
 TYPE: PRT
 ORGANISM: Homo sapiens
 PCT-US03-26780-3444

Query Match 9.6%; Score 80; DB 1; Length 1755;
 Best Local Similarity 43.9%; Pred. No. 1.1e+02;
 Matches 18; Conservative 1; Mismatches 22; Indels 0; Gaps 0;
 QY 31 GAATGAVVGVAGQLFGKSGRVAMAIGGAVLGGIGLIGSKIG 71
 DB 1268 GAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1308

RESULT 15
 US-09-674-546A-1669
 Sequence 1669, Application US/09674546A
 GENERAL INFORMATION:
 APPLICANT: Institute for Genomic Research
 TITLE OF INVENTION: Neisseria Meningitidis Antigens and Compositions
 FILE REFERENCE: CHIR-0334
 CURRENT APPLICATION NUMBER: US/09/674,546A
 CURRENT FILING DATE: 1999-04-30
 NUMBER OF SEQ ID NOS: 3264
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 1669
 LENGTH: 2599
 TYPE: PRT
 ORGANISM: Neisseria meningitidis
 US-09-674-546A-1669

Query Match 9.6%; Score 80; DB 5; Length 2599;
 Best Local Similarity 19.3%; Pred. No. 1.7e+02;
 Matches 40; Conservative 32; Mismatches 45; Indels 90; Gaps 8;
 QY 9 SLLIISVPLVGCANFSRQEV-----GAATGAVVGVAGQLFGK----- 49
 DB 2282 ASLGQWNAVL---AENQSRDYTWKGGIGRSLHGAAGGLTTGSLGGILAGGTSILAAPY 2338
 QY 50 -----SGRVAM-AIGGAVL---GGLIGSKIGQSMDDQD----- 78
 DB 2339 LDAAENLGPAGKAAVNAIGATGGSGGAVGANVDVNNRQLHPKEMALADKYAE 2398
 QY 79 -----KIKLNOSLEKVKAGQVTRWRNPDTGNSYSVEPV----- 111
 DB 2399 ALKREVEKREGKTISSQEAAMRIIR-QILRWVDKSGSDQGYTDSVISLIGMKGEDKALGY 2457
 QY 112 -----RTYQRYNKQERQYCRE 129
 DB 2458 TWDYRDYCARNPQTYNDPKLFEERYRQ 2484

Search completed: November 5, 2003, 20:25:58
 Job time : 10.2332 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 20:12:57 ; Search time 137.086 Seconds
(without alignments)
1068.647 Million cell updates/sec

Title: US-09-677-374-4

Perfect score: 832

Sequence: 1 MRGCLQGSSLIISVFLVGC.....IYGTACPODGRWQVISTEK 161

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 5728757 seqs, 909918778 residues

Total number of hits satisfying chosen parameters: 5728757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA Main:*

- 1: /cgn2_6/ptodata/1/paa/US06 COMB.pcp.*
- 2: /cgn2_6/ptodata/1/paa/US06 COMB.pcp.*
- 3: /cgn2_6/ptodata/1/paa/US07 COMB.pcp.*
- 4: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 5: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 6: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 7: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 8: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 9: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 10: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 11: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 12: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 13: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 14: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 15: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 16: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 17: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 18: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 19: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 20: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 21: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 22: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 23: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 24: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 25: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 26: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 27: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 28: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 29: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 30: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 31: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*
- 32: /cgn2_6/ptodata/1/paa/US08 COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	ID	Description
1	832	100.0	161 20	US-09-677-374-4 Sequence 4, Appli

2	832	100.0	161	28	US-10-261-446-4	Sequence 4, Appli
3	832	100.0	256	20	US-09-677-374-6	Sequence 6, Appli
4	832	100.0	256	28	US-10-261-446-6	Sequence 6, Appli
5	815	98.0	162	20	US-09-677-374-2	Sequence 2, Appli
6	815	98.0	162	28	US-10-261-446-2	Sequence 2, Appli
7	127.5	15.3	182	1	PCT-US02-12532-169	Sequence 169, App
8	127.5	15.3	182	27	US-10-127-032-169	Sequence 169, App
9	127.5	15.3	192	29	US-10-366-683-22817	Sequence 22817, A
10	127.5	15.3	192	30	US-10-419-128-22817	Sequence 22817, A
11	114.5	13.8	166	19	US-09-543-681A-4769	Sequence 4769, Ap
12	114.5	13.8	166	31	US-10-603-114-4769	Sequence 4769, Ap
13	112.5	13.5	155	22	US-09-791-537-85693	Sequence 85693, A
14	112	13.5	20	20	US-09-677-374-16	Sequence 16, Appl
15	112	13.5	20	28	US-10-261-446-16	Sequence 16, Appl
16	111.5	13.4	165	29	US-10-366-683-19701	Sequence 19701, A
17	111.5	13.4	165	30	US-10-419-128-19701	Sequence 19701, A
18	111	13.3	224	26	US-10-018-561-2	Sequence 2, Appli
19	111	13.3	224	26	US-10-018-561-4	Sequence 4, Appli
20	111	13.3	224	30	US-10-467-421-45	Sequence 45, Appl
21	111	13.3	231	19	US-09-540-236-3827	Sequence 3827, Ap
22	111	13.3	231	31	US-10-603-108-3827	Sequence 3827, Ap
23	111	13.3	231	32	US-60-128-476-3363	Sequence 3363, Ap
24	103	12.4	105	18	US-09-489-039A-10773	Sequence 10773, A
25	103	12.4	105	30	US-10-446-203-10773	Sequence 10773, A
26	101.5	12.2	83	21	US-09-739-449-12512	Sequence 12512, A
27	101.5	12.2	83	23	US-09-803-110-12512	Sequence 12512, A
28	101	12.1	154	22	US-09-791-537-117153	Sequence 117153,
29	101	12.1	155	28	US-10-274-586-319	Sequence 319, App
30	101	12.1	155	30	US-10-467-421-27	Sequence 27, Appl
31	100	12.0	155	18	US-09-489-039A-9682	Sequence 9682, Ap
32	100	12.0	155	30	US-10-446-203-9682	Sequence 9682, Ap
33	100	12.0	203	16	US-09-252-691-7664	Sequence 7664, Ap
34	100	12.0	203	16	US-09-252-691C-7664	Sequence 7664, Ap
35	100	12.0	203	30	US-10-417-886-7664	Sequence 7664, Ap
36	99.5	12.0	306	29	US-10-366-683-21754	Sequence 21754, A
37	99.5	12.0	306	30	US-10-419-128-21754	Sequence 21754, A
38	99.5	12.0	387	29	US-10-366-683-23334	Sequence 23334, A
39	99.5	12.0	387	30	US-10-419-128-23334	Sequence 23334, A
40	98	11.8	130	28	US-10-282-122A-68236	Sequence 68236, A
41	98	11.8	199	18	US-09-489-039A-7592	Sequence 7592, Ap
42	98	11.8	199	30	US-10-446-203-7592	Sequence 7592, Ap
43	95	11.4	175	22	US-09-758-472-9509	Sequence 9509, Ap
44	95	11.4	175	28	US-10-235-926-9509	Sequence 9509, Ap
45	95	11.4	231	19	US-09-540-209B-6649	Sequence 6649, Ap

ALIGNMENTS

RESULT 1
US-09-677-374-4
; Sequence 4, Application US/09677374
; GENERAL INFORMATION:
; APPLICANT: Burian, Jan
; APPLICANT: Kuzyk, Michael
; APPLICANT: Thornton, Julian
; APPLICANT: Kay, William
; TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY AGAINST
; FILE REFERENCE: IDCO1/60485/US
; CURRENT APPLICATION NUMBER: US/09/677,374
; CURRENT FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/154,437
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: NO 20004637
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: IE 2000/0752
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: GB 0022825.4
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: CL 2544-2000
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 20


```
; TITLE OF INVENTION: RICKETTSIAL DISEASES, AND ASSOCIATED PREVENTATIVE THERAPY
; FILE REFERENCE: 4616-64094
; CURRENT APPLICATION NUMBER: US/10/261,446
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US 09/677,374
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/154,437
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Piscirickettsia salmonis
US-10-261-446-2

Query Match      98.0%; Score 815; DB 20; Length 162;
Best Local Similarity 98.8%; Pred. No. 3.4e-81;
Matches 158; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2  RGCLOGSSLLIIISVFLVGCANFQSRQEVGAATGAVVGGVAGQLFGKSGRVSMAIGGAVL 61
      |||
Db      3  RGCLOGSSLLIIISVFLVGCANFQSRQEVGAATGAVVGGVAGQLFGKSGRVSMAIGGAVL 62
      |||

QY      62  GGLIGSKIGQSMDDQDKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTTYQRYNKQER 121
      |||
Db      63  GGLIGSKIGQSMDDQDKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTTYQRYNKQER 122
      |||

QY      122  RQCYCREFOQKAMIAQKQEIYGTACPDGRWQVISTEK 161
      |||
Db      123  RQCYCREFOQKAMIAQKQEIYGTACPDGRWQVISTEK 162
      |||

RESULT 7
PCT-US02-12532-169
; Sequence 169, Application PC/TUS0212532
; GENERAL INFORMATION:
; APPLICANT: University of Iowa Research Foundation, Inc., et al.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF
; TITLE OF INVENTION: BIOFILM FORMATION
; FILE REFERENCE: UIZ-070CPPEC
; CURRENT APPLICATION NUMBER: PCT/US02/12532
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/285,190
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: US 60/344,142
; PRIOR FILING DATE: 2001-10-24
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 169
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
PCT-US02-12532-169

Query Match      15.3%; Score 127.5; DB 1; Length 182;
Best Local Similarity 34.8%; Pred. No. 4.5e-05;
Matches 32; Conservative 13; Mismatches 28; Indels 19; Gaps 2;

QY      27  QEVGAATGAVVGGVAGQLFGKSGRVSMAIGGAVLGGILGSKIGQSMDDQDKI----- 79
      |||
Db      70  QIAGTAIGAVVGGLLGNQIGGGTGKKIATVAGAVGGVAGNKKVQEGMQBRDVTYTTTETRC 129
      |||

QY      80  -KLNQSLKVK-----KAGQVTRWRNP 99
      :|
Db      130  STVHDSSEKVGVDVKYMLDGRGQIRMERDP 161
      :|

RESULT 8
US-10-127-032-169
; Sequence 169, Application US/10127032
; GENERAL INFORMATION:
; APPLICANT: Whiteley, Marvin
```

```
; TITLE OF INVENTION: RICKETTSIAL DISEASES, AND ASSOCIATED PREVENTATIVE THERAPY
; FILE REFERENCE: 4616-64094
; CURRENT APPLICATION NUMBER: US/10/261,446
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US 09/677,374
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/154,437
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Piscirickettsia salmonis
US-09-677-374-2

Query Match      98.0%; Score 815; DB 20; Length 162;
Best Local Similarity 98.8%; Pred. No. 3.4e-81;
Matches 158; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      2  RGCLOGSSLLIIISVFLVGCANFQSRQEVGAATGAVVGGVAGQLFGKSGRVSMAIGGAVL 61
      |||
Db      3  RGCLOGSSLLIIISVFLVGCANFQSRQEVGAATGAVVGGVAGQLFGKSGRVSMAIGGAVL 62
      |||

QY      62  GGLIGSKIGQSMDDQDKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTTYQRYNKQER 121
      |||
Db      63  GGLIGSKIGQSMDDQDKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTTYQRYNKQER 122
      |||

QY      122  RQCYCREFOQKAMIAQKQEIYGTACPDGRWQVISTEK 161
      |||
Db      123  RQCYCREFOQKAMIAQKQEIYGTACPDGRWQVISTEK 162
      |||

RESULT 6
US-10-261-446-2
; Sequence 2, Application US/10261446
; GENERAL INFORMATION:
; APPLICANT: Kuzyk, Michael A.
; APPLICANT: Burian, Jan
; APPLICANT: Kay, William W.
; APPLICANT: Thornton, Julian C.
; TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY IN FISH AGAINST
```


RESULT 14
US-09-677-374-16
; Sequence 16, Application US/09677374
; GENERAL INFORMATION:

Qy 109 PVRTYQRYNKQERRRQYCRE 128
db 1 PVRTYQRYNKQERRRQYCRE 20

Search completed: November 5, 2003, 20:25:12
Job time : 137.086 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 20:13:32 ; Search time 9.17617 Seconds
(without alignments)
824,821 Million cell updates/sec

Title: US-09-677-374-4
Perfect score: 832
Sequence: 1 MRGCLQGSSLLIIISVFLVGC.....IYGTACPODGRWQVISTEK 161

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 248976

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_AA_New:*
1: /cgn2_6/prodata/2/paa/PCT_NEW_COMB.pcp.*
2: /cgn2_6/prodata/2/paa/US06_NEW_COMB.pcp.*
3: /cgn2_6/prodata/2/paa/US07_NEW_COMB.pcp.*
4: /cgn2_6/prodata/2/paa/US08_NEW_COMB.pcp.*
5: /cgn2_6/prodata/2/paa/US09_NEW_COMB.pcp.*
6: /cgn2_6/prodata/2/paa/US10_NEW_COMB.pcp.*
7: /cgn2_6/prodata/2/paa/US60_NEW_COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	832	100.0	161	6	US-10-261-445B-4
2	832	100.0	256	6	US-10-261-445B-6
3	815	98.0	162	6	US-10-241-602B-6
4	815	98.0	162	6	US-10-261-445B-2
5	112	13.5	20	6	US-10-261-445B-16
6	102.5	12.3	223	5	US-09-581-286A-434
7	102.5	12.3	230	5	US-09-581-286A-309
8	94.5	11.4	106	5	US-09-897-516A-5133
9	86.5	10.4	99	6	US-10-389-647-605
10	83	10.0	423	1	PCT/US02/39429-8
11	82.5	9.9	734	6	US-10-425-114A-49440
12	79.5	9.6	1117	1	PCT-US03-21510-114
13	79	9.5	593	6	US-10-435-696-50
14	78.5	9.5	1261	6	US-10-679-063-16542
15	78.5	9.4	261	6	US-10-689-006-24
16	78	9.4	403	5	US-09-897-516A-7766
17	78	9.4	438	1	PCT-US02-39429-88
18	78	9.4	515	6	US-10-679-063-22745
19	78	9.4	517	6	US-10-679-063-22744
20	77	9.3	1755	1	PCT-US03-26780-3444
21	77	9.3	2599	5	US-09-674-546A-1669
22	73	8.8	254	1	PCT-US03-10856A-5
23	73	8.8	293	1	PCT-US03-28227-2863
24	73	8.8	800	6	US-10-296-115-1229
25	73	8.8	820	1	PCT-US03-28227-5403
26	73	8.8	871	1	PCT-US03-31442-3

27	73	8.8	2322	1	PCT-US03-26780-2711	Sequence 2711, Ap
28	72.5	8.7	260	6	US-10-435-614-19	Sequence 19, Appl
29	72.5	8.7	348	6	US-10-679-063-7732	Sequence 7732, Ap
30	72	8.7	429	6	US-10-425-114A-64061	Sequence 64061, A
31	72	8.7	596	6	US-10-425-114A-52050	Sequence 52050, A
32	72	8.7	1064	6	US-10-418-861B-77	Sequence 77, Appl
33	71.5	8.6	181	7	US-10-425-114A-59401	Sequence 59401, A
34	71.5	8.6	315	6	US-10-499-964-418	Sequence 418, Appl
35	71.5	8.6	456	5	US-09-830-321A-3	Sequence 3, Appli
36	71.5	8.6	456	5	US-09-976-858-128	Sequence 128, App
37	71.5	8.6	690	1	PCT-US03-28227-4797	Sequence 4797, Ap
38	71.5	8.6	733	5	US-09-743-818A-4	Sequence 4, Appli
39	71.5	8.6	798	1	PCT-US03-18234-38	Sequence 38, Appl
40	71.5	8.6	798	7	US-60-487-610-2779	Sequence 2779, Ap
41	71.5	8.6	838	1	PCT-US03-28227-4796	Sequence 4796, Ap
42	71	8.5	89	6	US-10-472-928-932	Sequence 932, App
43	71	8.5	224	6	US-10-425-114A-43688	Sequence 43688, A
44	71	8.5	230	6	US-10-425-114A-72589	Sequence 72589, A
45	71	8.5	260	6	US-10-435-614-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1

US-10-261-445B-4

; Sequence 4, Application US/10261445B

; GENERAL INFORMATION:

; APPLICANT: Kuzyk, Michael A.

; APPLICANT: Burian, Jan

; APPLICANT: Thornton, Julian C.

; TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY IN FISH AGAINST

; TITLE OF INVENTION: RICKETTSIAL DISEASES, AND ASSOCIATED PREVENTATIVE THERAPY

; FILE REFERENCE: 4616-64095

; CURRENT APPLICATION NUMBER: US/10/261,445B

; CURRENT FILING DATE: 2002-09-30

; PRIOR APPLICATION NUMBER: US 09/677,374

; PRIOR FILING DATE: 2000-09-15

; PRIOR APPLICATION NUMBER: US 60/154,437

; PRIOR FILING DATE: 1999-09-17

; NUMBER OF SEQ ID NOS: 20

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 4

; LENGTH: 161

; TYPE: PRT

; ORGANISM: Piscirickettsia salmonis

US-10-261-445B-4

Query Match 100.0%; Score 832; DB 6; Length 161;
Best Local Similarity 100.0%; Pred. No. 3.2e-66;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRGCLQGSSLLIIISVFLVGCQNFQRQEVGAATGAVGVGAGQLFGKSGRVSMAIGGAV 60
|||||
Db 1 MRGCLQGSSLLIIISVFLVGCQNFQRQEVGAATGAVGVGAGQLFGKSGRVSMAIGGAV 60
|||||

QY 61 LGLLIGSKIGQSMDDQDKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEVPRTYQRYNKQE 120
|||||
Db 61 LGLLIGSKIGQSMDDQDKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEVPRTYQRYNKQE 120
|||||

QY 121 RQOQYCRBFQOQKAMTAGQKQEIYGTACPODGRWQVISTEK 161
|||||

Db 121 RQOQYCRBFQOQKAMTAGQKQEIYGTACPODGRWQVISTEK 161
|||||

RESULT 2

US-10-261-445B-5

; Sequence 6, Application US/10261445B

; GENERAL INFORMATION:

; APPLICANT: Kuzyk, Michael A.

; APPLICANT: Burian, Jan

; APPLICANT: Kay, William W.

APPLICANT: Thornton, Julian C.
TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY IN FISH AGAINST
TITLE OF INVENTION: RICKETTSIAL DISEASES, AND ASSOCIATED PREVENTATIVE THERAPY
FILE REFERENCE: 4616-64095
CURRENT APPLICATION NUMBER: US/10/261,445B
PRIOR FILING DATE: 2002-09-30
PRIOR APPLICATION NUMBER: US 09/677,374
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/154,437
PRIOR FILING DATE: 1999-09-17
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.0
SEQ ID NO 6
LENGTH: 256
TYPE: PRT
ORGANISM: Piscirickettsia salmonis
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (-95)...(-1)
US-10-261-445B-6

Query Match 100.0%; Score 832; DB 6; Length 256;
Best Local Similarity 100.0%; Pred. No. 5.5e-66;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRCLGSSLIISVFLVGCANFSGQEVGAATGAVVGGVAGOLFCKGSGRVSMAIGGAV 60
DB 96 MRCLGSSLIISVFLVGCANFSGQEVGAATGAVVGGVAGOLFCKGSGRVSMAIGGAV 155
QY 61 LGLGSKIGSQSMDQDDKIKLNQSLKVKAGQVTRWRNPDGTGNSYSVEPVRTYQRYNKB 120
DB 156 LGLGSKIGSQSMDQDDKIKLNQSLKVKAGQVTRWRNPDGTGNSYSVEPVRTYQRYNKB 215
QY 121 RQOYCREFQOKAMIAQKQEIYGTACPDGGRWQVISTEK 161
DB 216 RQOYCREFQOKAMIAQKQEIYGTACPDGGRWQVISTEK 256

RESULT 3
US-10-241-602B-6
Sequence 6, Application US/10241602B
GENERAL INFORMATION:
APPLICANT: Simard, Nathalie
APPLICANT: Brouwers, Huub
APPLICANT: Jones, Simon
APPLICANT: Griffiths, Steve
APPLICANT: Valenzuela, Pablo
APPLICANT: Burzio, Luis
TITLE OF INVENTION: Sequences from Piscirickettsia salmonis
FILE REFERENCE: H-32319A
CURRENT APPLICATION NUMBER: US/10/241,602B
CURRENT FILING DATE: 2002-09-11
PRIOR APPLICATION NUMBER: PCT/GB01/01055
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: GB0005838.8
PRIOR FILING DATE: 2000-03-11
PRIOR APPLICATION NUMBER: GB0016080.4
PRIOR FILING DATE: 2000-07-01
PRIOR APPLICATION NUMBER: GB0016082.0
PRIOR FILING DATE: 2000-07-01
PRIOR APPLICATION NUMBER: GB0018599.1
PRIOR FILING DATE: 2000-07-29
NUMBER OF SEQ ID NOS: 34
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 162
TYPE: PRT
ORGANISM: Piscirickettsia salmonis
US-10-241-602B-6

Query Match 99.0%; Score 815; DB 6; Length 162;
Best Local Similarity 98.8%; Pred. No. 1e-64;
Matches 158; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 RGCLOGSSLIISVFLVGCANFSGQEVGAATGAVVGGVAGOLFCKGSGRVSMAIGGAVL 61
DB 3 RGCLOGSSLIISVFLVGCANFSGQEVGAATGAVVGGVAGOLFCKGSGRVSMAIGGAVL 62
QY 62 GGLGSKIGSQSMDQDDKIKLNQSLKVKAGQVTRWRNPDGTGNSYSVEPVRTYQRYNKB 121
DB 63 GGLGSKIGSQSMDQDDKIKLNQSLKVKAGQVTRWRNPDGTGNSYSVEPVRTYQRYNKB 122
QY 122 RQOYCREFQOKAMIAQKQEIYGTACPDGGRWQVISTEK 161
DB 123 RQOYCREFQOKAMIAQKQEIYGTACPDGGRWQVISTEK 162

RESULT 4
US-10-261-445B-2
Sequence 2, Application US/10261445B
GENERAL INFORMATION:
APPLICANT: Kuzyk, Michael A.
APPLICANT: Burian, Jan
APPLICANT: Kay, William W.
APPLICANT: Thornton, Julian C.
TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY IN FISH AGAINST
TITLE OF INVENTION: RICKETTSIAL DISEASES, AND ASSOCIATED PREVENTATIVE THERAPY
FILE REFERENCE: 4616-64095
CURRENT APPLICATION NUMBER: US/10/261,445B
CURRENT FILING DATE: 2002-09-30
PRIOR APPLICATION NUMBER: US 09/677,374
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/154,437
PRIOR FILING DATE: 1999-09-17
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 162
TYPE: PRT
ORGANISM: Piscirickettsia salmonis
US-10-261-445B-2

Query Match 98.0%; Score 815; DB 6; Length 162;
Best Local Similarity 98.8%; Pred. No. 1e-64;
Matches 158; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 2 RGCLOGSSLIISVFLVGCANFSGQEVGAATGAVVGGVAGOLFCKGSGRVSMAIGGAVL 61
DB 3 RGCLOGSSLIISVFLVGCANFSGQEVGAATGAVVGGVAGOLFCKGSGRVSMAIGGAVL 62
QY 62 GGLGSKIGSQSMDQDDKIKLNQSLKVKAGQVTRWRNPDGTGNSYSVEPVRTYQRYNKB 121
DB 63 GGLGSKIGSQSMDQDDKIKLNQSLKVKAGQVTRWRNPDGTGNSYSVEPVRTYQRYNKB 122
QY 122 RQOYCREFQOKAMIAQKQEIYGTACPDGGRWQVISTEK 161
DB 123 RQOYCREFQOKAMIAQKQEIYGTACPDGGRWQVISTEK 162

RESULT 5
US-10-261-445B-16
Sequence 16, Application US/10261445B
GENERAL INFORMATION:
APPLICANT: Kuzyk, Michael A.
APPLICANT: Burian, Jan
APPLICANT: Kay, William W.
APPLICANT: Thornton, Julian C.
TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY IN FISH AGAINST
TITLE OF INVENTION: RICKETTSIAL DISEASES, AND ASSOCIATED PREVENTATIVE THERAPY
FILE REFERENCE: 4616-64095
CURRENT APPLICATION NUMBER: US/10/261,445B
CURRENT FILING DATE: 2002-09-30
PRIOR APPLICATION NUMBER: US 09/677,374
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/154,437
PRIOR FILING DATE: 1999-09-17

APPLICANT: SPILLUONOV, SERGEI
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof

```

: FILE REFERENCE: 38-21(51847) B
: CURRENT APPLICATION NUMBER: US/09/897,516A
: CURRENT FILING DATE: 2001-06-29
: PRIOR APPLICATION NUMBER: US 6/0215,161
: PRIOR FILING DATE: 2000-06-30
: NUMBER OF SEQ ID NOS: 8415
: SEQ ID NO 5133
: LENGTH: 106
: TYPE: PRT
: ORGANISM: Xenorhabdus sp.
US-09-897-516A-5133

```

Query Match 11.4%; Score 94.5; DB 5; Length 106;
Best Local Similarity 35.1%; Pred. No. 0.16;
Matches 27; Conservative 9; Mismatches 32; Indels 9; Gaps 2;

```

RESULT 9
US-10-389-647-605
; Sequence 605, Application US/10389647
; GENERAL INFORMATION:
; APPLICANT: GREENBERG, E. Peter
; APPLICANT: SCHUSTER, Martin
; APPLICANT: LOSTROH, Candi
; TITLE OF INVENTION: QUORUM SENSING SIGNALING IN BACTERIA
; FILE REFERENCE: UTZ-038CP
; CURRENT APPLICATION NUMBER: US/10/389,647
; CURRENT FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: 09/653730
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/153022
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 710
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 605
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-389-647-605

```

Query Match	10.4%	Score 86.5;	DB 6;	Length 99;
Best Local Similarity	34.3%	Pred. No. 0.74;		
Matches	24;	Conservative	8;	Mismatches 21; Indels 17; Gaps 3;
Qy	4	CLOGSSLIILISVFLVGCQAQNF	SROEVGAATCAVVGGVAGOLF	KGSGRVSMAIG-----57
Db	11	CVSSGTL-----SGMIVGAVD	-----CAATGMAIGKRWGGAGG	FGFGLSQLVGLIIVPTA 60
Qy	58	-GAVLGSLIG	66	
Db	61	MGAIAAGGTVG	70	

RESULT 10
PCT/US02/39429-8
; Sequence 8, Application PC/TUS0239429
; GENERAL INFORMATION:
; APPLICANT: NeoRx Corporation
; APPLICANT: Goshorn, Stephen Charles
; APPLICANT: Graves, Scott Stoll
; APPLICANT: Schultz, Joanne Elaine
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James Allen
; APPLICANT: Reno, John M.
; APPLICANT: Dearstyne, Erica A.

```

; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 690022.54701PC
; CURRENT APPLICATION NUMBER: PCT/US02/39429
; CURRENT FILING DATE: 2002-12-06
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 423
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Predicted amino acid sequence of B9E9 s
PCT/US02/19429-8

```

Query Match	10.0%	Score 83;	DB 1;	Length 423;
Best Local Similarity	21.7%;	Pred. No. 8.7;		
Matches	33;	Conservative 24;	Mismatches 67;	Indels 28; Gaps 5
Qy	30	GAATGAVVGGVAGOLF	GGSGRVSMTGGAVL	GLGSLKIQSQSDQDKIKLQ----- 83
Db	134	GGSGGGGGGGSG---	GGGSSDIVLSQSPAIL	SLASPGKVTMTCRASSSVSMHWYQKP 190
Qy	84	-----SLEKVKAGQVTR	WRNPDTGNSVPEPVTY	QRYNKKRRQQYCRFP--QQK 132
Db	191	GSPEKPMIYATSNLASG	VPARFSGSGGTSVSL---	TISVEAEDAAITYCQQWISNPP 246
Qy	133	AMTAGQKEI-----	YGTACPOPQDGRWQVISTE	160
Db	247	TFGAGTKLELKSSGGSG	ADPFSKQSKAQVSAAE	278

```

RESULT 11
US-10-425-114A-49440
; Sequence 49440, Application US/10425114A
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114A
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 49440
; LENGTH: 734
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700991377_FLI pep
US-10-425-114A-49440

```

Query Match	9.9%	Score 82.5;	DB 6;	Length 734;
Best Local Similarity	25.5%;	Pred. No. 19;		
Matches	35;	Conservative 25;	Mismatches 48;	Indels 29; Gaps 6;
Qy	25	SREVGNAATCAVVGGA--CQLFKGSGRVS-----	MAIGGAVLG---CLISKI	69
Db	77	SKNEVNGVAFADGNGVTEHGEVWGSKNEAVNNEVAIVDGNVGAVYGGVEHGSKNGAVNNEV	136	
Qy	70	-----GQSMDDQDKIKLNGSLKLVKAGQVTRWRNPDTGNSYSVEPVRTTYQRYNKQERRQ	123	
Db	137	VIADGNGVTEQGDHCLKNETVNVVA-----	NADEGNSGAVECQTYKRRKHAKSSS	189
Qy	124	QY-CREFOQKAMTAGQK	139	
Db	190	EPKVOENSRKHMGAASO	206	

```

RESULT 12
PCT-US03-21510-114
; Sequence 114, Application PC/TUS0321510
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: MP21S AS MODIFIERS OF THE P21 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX03-046C-PC
; CURRENT APPLICATION NUMBER: PCT/US03/21510
; CURRENT FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: US 60/394,795
; PRIOR FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/401,739
; PRIOR FILING DATE: 2002-08-07
; PRIOR APPLICATION NUMBER: US 60/411,010
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/437,158
; PRIOR FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 114
; LENGTH: 1117
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-21510-114

Query Match          9.6%; Score 79.5; DB 1; Length 1117;
Best Local Similarity 21.2%; Pred. No. 57;
Matches 41; Conservative 25; Mismatches 60; Indels 67; Gaps 7;

QY 23 NPSRQEVGAATGAVVG-GVAGQ-----LFGKSGRVSMAIGAVLGLGSKIGQSMDDQ 76
DB 381 NISTSNIPSAAGSVSGVGTSGVNVNLLSGMNGITSSAAVSSVPNAAAGTGGVSQQ 440
QY 77 DK-----IKNOSLEKVKAG-----
DB 441 CQQPTVNTSRFRVVKLSDSSSEPFKGRWTCBFEYKENAVPATEGLINKVETVKQNP 500
QY 92 QVTRWRNPDTGNSYSEVPTTYQRYNK-----QERRQOYCREFOQKAMTAG 137
DB 501 EVTSRESTSGSVS-SVSTLSHTSYESVSGEMGAPTVVVQQQQQQQQQQQQPALQGV 559
QY 138 QKQEI-YGTACPQ 149
DB 560 TLQWDFGSGTGPQ 572

RESULT 13
US-10-435-696-50
; Sequence 50, Application US/10435696
; GENERAL INFORMATION:
; APPLICANT: Wirtz, Ralph
; APPLICANT: Munnes, Marc
; APPLICANT: Kallabis, Harald
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE PREDICTION, DIAGNOSIS, PROGNOSIS
; PREVENTION AND TREATMENT OF MALIGNANT NEOPLASIA
; FILE REFERENCE: Lea 36 108
; CURRENT APPLICATION NUMBER: US/10/435,696
; CURRENT FILING DATE: 2003-05-09
; PRIOR APPLICATION NUMBER: EP03003112.4
; PRIOR FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: EP02010291.9
; PRIOR FILING DATE: 2002-05-21
; NUMBER OF SEQ ID NOS: 314
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 50
; LENGTH: 593
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-435-696-50

Query Match          9.5%; Score 79; DB 6; Length 593;
Best Local Similarity 28.0%; Pred. No. 29;
Matches 33; Conservative 13; Mismatches 46; Indels 26; Gaps 7;

```

```

QY 23 NPSRQEVGAATGAVVG--VAGOLFQKGS-----GRVSMIAIGAVLGLGSKIGQSMDDQ 76
DB 88 SYGSSSGSVGSGFGGNGFGGSGFGGSGFGGSGFGGSGFGGSGFGGSGFGGSGFGGSG 145
QY 77 DK-----KLNOSLEKVK-----QVTRWRNPDTGNSYSEVPTTYQRYNK 118
DB 146 EKVTMQNLNDRLASYLQKVRALBESNTELEGIKIEWEKH-GNSHQGEP-RDYSKYK 201

RESULT 14
US-10-679-063-16542
; Sequence 16542, Application US/10679063
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
; FILE REFERENCE: 38-15(52054)B
; CURRENT APPLICATION NUMBER: US/10/679,063
; CURRENT FILING DATE: 2003-10-02
; PRIOR APPLICATION NUMBER: 60/415,758
; PRIOR FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 27373
; SEQ ID NO 16542
; LENGTH: 1261
; TYPE: PRT
; ORGANISM: Magnetococcus sp. MC-1
US-10-679-063-16542

Query Match          9.5%; Score 79; DB 6; Length 1261;
Best Local Similarity 24.8%; Pred. No. 73;
Matches 30; Conservative 17; Mismatches 32; Indels 42; Gaps 6;

QY 30 GAATGAVVG-----GVAGOLFQKGS-----RVSMIAIG-----AVLGLI 65
DB 200 GOLGAVASLVASHIRDLGLANLIIGKVAGPPLMLRVGVKGAPELLQAFQAVVDGI- 258
QY 66 GSKIGQSMDDQDKIKNOSLEKVKAGQVTRWRNPDTGNSYSEVPTTYQRYNKERRQY 125
DB 259 -----SQEDHIRTKWHPVLAGELDANR-----LSLNPAQ--QAVLQEKKRISY 301
QY 126 C 126
DB 302 C 302

RESULT 15
US-10-689-006-24
; Sequence 24, Application US/10689006
; GENERAL INFORMATION:
; APPLICANT: Vanderbilt University
; APPLICANT: Hallahan, Dennis E
; APPLICANT: Mernaugh, Raymond
; TITLE OF INVENTION: PHAGE ANTIBODIES TO RADIATION-INDUCIBLE NEOANTIGENS
; FILE REFERENCE: 1242/72
; CURRENT APPLICATION NUMBER: US/10/689,006
; CURRENT FILING DATE: 2003-10-20
; PRIOR APPLICATION NUMBER: US 09/914,605
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: US 10/259,087
; PRIOR FILING DATE: 2002-09-27
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 24
; LENGTH: 261
; TYPE: PRT
; ORGANISM: artificial
; FEATURE:
; OTHER INFORMATION: nucleic acid encoding scFv antibody 10A
US-10-689-006-24

Query Match          9.4%; Score 78.5; DB 6; Length 261;
Best Local Similarity 19.3%; Pred. No. 12;
Matches 26; Conservative 19; Mismatches 65; Indels 25; Gaps 3;

```

```
QY 38 GGVAGOLFQKGRVSWAICGAVLGLGKIGKIQSMDQODKIKINOSLEK----- 87
Db 124 GSGGGGGGGSDIELTQSPAIMSATLGEKVTMSCRASSNVKYWYQKSGASPKLWI 183
QY 88 -----VKAGOVTRWRNPDTGNSYSVEPVRTYORYNKOERRQOYCREFOOKAMIAGQ----- 138
Db 184 YTSNLAGVVPARFSGSGGTSYSL----TISVEAEDAATYYCQOFTSSPYTFGSGTKL 239
QY 139 --KQEIYGTACPOPD 151
Db 240 EIKRAAAGAPVPYPD 254
```

Search completed: November 5, 2003, 20:25:58
Job time : 9.17617 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 20:12:57 ; Search time 217.976 Seconds
(without alignments)
1068.647 Million cell updates/sec

Title: US-09-677-374-6
Perfect score: 1335
Sequence: 1 MSVEFYNGSKAQTSITPI.....LYGTACPODGRQWVISTEK 256

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 5728757 seqs, 909918778 residues

Total number of hits satisfying chosen parameters: 5728757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents AA Main:*

1:	/cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
2:	/cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3:	/cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4:	/cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5:	/cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6:	/cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7:	/cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8:	/cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9:	/cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10:	/cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11:	/cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12:	/cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13:	/cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14:	/cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15:	/cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16:	/cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17:	/cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18:	/cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19:	/cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20:	/cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21:	/cgn2_6/ptodata/1/paa/US097A_COMB.pep.*
22:	/cgn2_6/ptodata/1/paa/US097B_COMB.pep.*
23:	/cgn2_6/ptodata/1/paa/US098_COMB.pep.*
24:	/cgn2_6/ptodata/1/paa/US099A_COMB.pep.*
25:	/cgn2_6/ptodata/1/paa/US099B_COMB.pep.*
26:	/cgn2_6/ptodata/1/paa/US100_COMB.pep.*
27:	/cgn2_6/ptodata/1/paa/US101_COMB.pep.*
28:	/cgn2_6/ptodata/1/paa/US102_COMB.pep.*
29:	/cgn2_6/ptodata/1/paa/US103_COMB.pep.*
30:	/cgn2_6/ptodata/1/paa/US104_COMB.pep.*
31:	/cgn2_6/ptodata/1/paa/US106_COMB.pep.*
32:	/cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	ID	Description
1	1335	100.0	256	20 US-09-677-374-6 Sequence 6, Appli

2	1335	100.0	256	28	US-10-261-446-6	Sequence 6, Appli
3	832	62.3	161	20	US-09-677-374-4	Sequence 4, Appli
4	832	62.3	161	28	US-10-261-446-4	Sequence 4, Appli
5	815	61.0	162	20	US-09-677-374-2	Sequence 2, Appli
6	815	61.0	162	28	US-10-261-446-2	Sequence 2, Appli
7	487	36.5	190	1	PCT-US000-00708-4	Sequence 4, Appli
8	487	36.5	326	1	PCT-US000-00708-10	Sequence 10, Appli
9	487	36.5	340	1	PCT-US000-00708-6	Sequence 6, Appli
10	483	36.2	163	1	PCT-US000-00708-2	Sequence 2, Appli
11	483	36.2	426	1	PCT-US000-00708-8	Sequence 8, Appli
12	480.5	36.0	382	23	US-09-808-898-22	Sequence 22, Appli
13	477	35.7	162	7	US-08-330-394-2	Sequence 2, Appli
14	477	35.7	162	13	US-08-962-588-2	Sequence 2, Appli
15	466.5	34.9	163	30	US-10-460-524-2	Sequence 2, Appli
16	466.5	34.9	163	32	US-60-388-059-2	Sequence 2, Appli
17	249	18.7	586	22	US-09-791-537-105096	Sequence 105096
18	245	18.4	167	12	US-08-817-940-9	Sequence 9, Appli
19	245	18.4	476	29	US-10-372-054-4	Sequence 4, Appli
20	245	18.4	493	15	US-09-198-955-12	Sequence 12, Appli
21	245	18.4	493	26	US-10-072-152-12	Sequence 12, Appli
22	245	18.4	531	9	US-08-559-968-29	Sequence 29, Appli
23	245	18.4	599	25	US-09-955-555A-29	Sequence 29, Appli
24	245	18.4	772	22	US-03-791-537-65163	Sequence 65163, A
25	245	18.4	1853	22	US-09-791-537-19761	Sequence 19761, A
26	245	18.4	1854	22	US-03-791-537-124459	Sequence 124459,
27	224	16.8	1546	22	US-09-791-537-6661	Sequence 6661, Ap
28	223.5	16.7	1483	28	US-10-282-122A-51483	Sequence 51483, A
29	193	14.5	887	22	US-09-791-537-17654	Sequence 17654, A
30	188.5	14.1	879	22	US-09-791-537-10882	Sequence 10882, A
31	188	14.1	946	22	US-09-791-537-74713	Sequence 74713, A
32	188	14.1	986	22	US-09-791-537-30993	Sequence 30993, A
33	169.5	12.7	321	22	US-09-791-537-84844	Sequence 84844, A
34	162.5	12.2	914	22	US-09-791-537-106665	Sequence 106665,
35	152.5	11.4	997	22	US-09-791-537-48723	Sequence 48723, A
36	149.5	11.2	1426	13	US-08-932-571-43	Sequence 43, Appli
37	149.5	11.2	1426	21	US-09-723-539B-43	Sequence 43, Appli
38	149.5	11.2	1426	21	US-09-723-539B-43	Sequence 43, Appli
39	149.5	11.2	1426	22	US-09-791-537-597	Sequence 597, App
40	148.5	11.1	616	21	US-09-723-539B-47	Sequence 47, Appli
41	148.5	11.1	1751	13	US-08-932-571-44	Sequence 44, Appli
42	148.5	11.1	1751	21	US-09-723-539B-44	Sequence 44, Appli
43	148.5	11.1	1751	21	US-09-723-539B-44	Sequence 44, Appli
44	147	11.0	915	22	US-09-791-537-74930	Sequence 74930, A
45	143.5	10.7	1331	22	US-09-791-537-81968	Sequence 81968, A

ALIGNMENTS

RESULT 1
US-09-677-374-6
; Sequence 6, Application US/09677374
; GENERAL INFORMATION:
; APPLICANT: Burian, Jan
; APPLICANT: Kuzyk, Michael
; APPLICANT: Thornton, Julian
; APPLICANT: Kay, William
; TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY AGAINST
; FILE REFERENCE: IDCO1/60485/US
; CURRENT APPLICATION NUMBER: US/09/677,374
; CURRENT FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/154,437
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: NO 20004637
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: IE 2000/0752
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: GB 0022825.4
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: CL 2544-2000
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 20

```
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 6
/ LENGTH: 256
/ TYPE: PRT
/ ORGANISM: Piscirickettsia salmonis
/ FEATURE:
/ NAME/KEY: SIGNAL
/ LOCATION: (-95)..(-1)
US-09-677-374-6

Query Match      100.0%; Score 1335; DB 20; Length 256;
Best Local Similarity 100.0%; Pred. No. 7.1e-131;
Matches 256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSVEFNSNKSAAQTNSITPIIKITNTSDSLNLDVKNVRYTSDGTQGTQWCDHAGAL 60
    |||||||
Db 1 MSVEFNSNKSAAQTNSITPIIKITNTSDSLNLDVKNVRYTSDGTQGTQWCDHAGAL 60
    |||||||

QY 61 LGNSYVDNTSKVTANFVKETASPTSTYDYLDPFSHMRGCLQGSSLLIISVFLVGCAQNF 120
    |||||||
Db 61 LGNSYVDNTSKVTANFVKETASPTSTYDYLDPFSHMRGCLQGSSLLIISVFLVGCAQNF 120
    |||||||

QY 121 ROEVGAATGAVGVAGQLFGKSGSRVSMAGVGLGLIGSKIGQSMDDQDKIKLNQSL 180
    |||||||
Db 121 ROEVGAATGAVGVAGQLFGKSGSRVSMAGVGLGLIGSKIGQSMDDQDKIKLNQSL 180
    |||||||

QY 121 ROEVGAATGAVGVAGQLFGKSGSRVSMAGVGLGLIGSKIGQSMDDQDKIKLNQSL 180
    |||||||
Db 121 ROEVGAATGAVGVAGQLFGKSGSRVSMAGVGLGLIGSKIGQSMDDQDKIKLNQSL 180
    |||||||

QY 181 EKVAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQERRQQYCRFQOKAMIAGQKOEIYGT 240
    |||||||
Db 181 EKVAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQERRQQYCRFQOKAMIAGQKOEIYGT 240
    |||||||

QY 181 EKVAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQERRQQYCRFQOKAMIAGQKOEIYGT 240
    |||||||
Db 181 EKVAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQERRQQYCRFQOKAMIAGQKOEIYGT 240
    |||||||

QY 241 ACQPQDGRWQVISTEK 256
    |||||||
Db 241 ACQPQDGRWQVISTEK 256
    |||||||

US-10-261-446-6
/ Sequence 6, Application US/10261446
/ GENERAL INFORMATION:
/ APPLICANT: Kuzyk, Michael A.
/ APPLICANT: Burian, Jan
/ APPLICANT: Thornton, Julian C.
/ TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY IN FISH AGAINST
/ TITLE OF INVENTION: RICKETTSIAL DISEASES, AND ASSOCIATED PREVENTATIVE THERAPY
/ FILE REFERENCE: 4616-64094
/ CURRENT APPLICATION NUMBER: US/10/261,446
/ CURRENT FILING DATE: 2002-09-30
/ PRIOR APPLICATION NUMBER: US 09/677,374
/ PRIOR FILING DATE: 2000-09-15
/ PRIOR APPLICATION NUMBER: US 60/154,437
/ PRIOR FILING DATE: 1999-09-17
/ NUMBER OF SEQ ID NOS: 20
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 6
/ LENGTH: 256
/ TYPE: PRT
/ ORGANISM: Piscirickettsia salmonis
US-10-261-446-6

Query Match      100.0%; Score 1335; DB 28; Length 256;
Best Local Similarity 100.0%; Pred. No. 7.1e-131;
Matches 256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSVEFNSNKSAAQTNSITPIIKITNTSDSLNLDVKNVRYTSDGTQGTQWCDHAGAL 60
    |||||||
Db 1 MSVEFNSNKSAAQTNSITPIIKITNTSDSLNLDVKNVRYTSDGTQGTQWCDHAGAL 60
    |||||||

QY 61 LGNSYVDNTSKVTANFVKETASPTSTYDYLDPFSHMRGCLQGSSLLIISVFLVGCAQNF 120
    |||||||
```


FILE REFERENCE: 4616-64094
CURRENT APPLICATION NUMBER: US/10/261,446
CURRENT FILING DATE: 2002-09-30
PRIOR APPLICATION NUMBER: US 09/677,374
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/154,437
PRIOR FILING DATE: 1999-09-17
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4
LENGTH: 161
TYPE: PRT
ORGANISM: Piscirickettsia salmonis
US-10-261-446-4

Query Match 62.3%; Score 832; DB 28; Length 161;
Best Local Similarity 100.0%; Pred. No. 2.4e-78;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 96 MRGCLQGSSLLIIISVFLVGCNQFNRQEVGAATGAVVGGVAGQLFGKSGRVSMAIGGAV 155
DB 1 MRGCLQGSSLLIIISVFLVGCNQFNRQEVGAATGAVVGGVAGQLFGKSGRVSMAIGGAV 60
QY 156 LGGLGSKIGQSMDDQDKIKLQNSLEKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKOE 215
DB 61 LGGLGSKIGQSMDDQDKIKLQNSLEKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKOE 120
QY 216 RQOYCREPQOKAMTAGQKOEIYGTACPODGRWQVISTEK 256
DB 121 RQOYCREPQOKAMTAGQKOEIYGTACPODGRWQVISTEK 161

RESULT 5
US-09-677-374-2
Sequence 2, Application US/09677374
GENERAL INFORMATION:
APPLICANT: Burian, Jan
APPLICANT: Kuzyk, Michael
APPLICANT: Thornton, Julian
APPLICANT: Kay, William
TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY AGAINST
TITLE OF INVENTION: RICKETTSIAL DISEASES, AND ASSOCIATED PREVENTATIVE THERAPY
FILE REFERENCE: IDC01/60485/US
CURRENT APPLICATION NUMBER: US/09/677,374
CURRENT FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/154,437
PRIOR FILING DATE: 1999-09-17
PRIOR APPLICATION NUMBER: NO 20004637
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: IE 2000/0752
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: GB 00282825.4
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: CL 2544-2000
PRIOR FILING DATE: 2000-09-15
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 162
TYPE: PRT
ORGANISM: Piscirickettsia salmonis
US-09-677-374-2

Query Match 61.0%; Score 815; DB 20; Length 162;
Best Local Similarity 98.8%; Pred. No. 1.5e-76;
Matches 158; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 97 RGCLQGSSLLIIISVFLVGCNQFNRQEVGAATGAVVGGVAGQLFGKSGRVSMAIGGAVL 156
DB 3 RGCLQGSSLLIIISVFLVGCNQFNRQEVGAATGAVVGGVAGQLFGKSGRVSMAIGGAVL 62
QY 157 GGLGSKIGQSMDDQDKIKLQNSLEKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKOE 216

DB 63 GGLGSKIGQSMDDQDKIKLQNSLEKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKOE 122
QY 217 RQOYCREPQOKAMTAGQKOEIYGTACPODGRWQVISTEK 256
DB 123 RQOYCREPQOKAMTAGQKOEIYGTACPODGRWQVISTEK 162

RESULT 6
US-10-261-446-2
Sequence 2, Application US/10261446
GENERAL INFORMATION:
APPLICANT: Kuzyk, Michael A.
APPLICANT: Burian, Jan
APPLICANT: Kay, William W.
APPLICANT: Thornton, Julian C.
TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY IN FISH AGAINST
TITLE OF INVENTION: RICKETTSIAL DISEASES, AND ASSOCIATED PREVENTATIVE THERAPY
FILE REFERENCE: 4616-64094
CURRENT APPLICATION NUMBER: US/10/261,446
CURRENT FILING DATE: 2002-09-30
PRIOR APPLICATION NUMBER: US 09/677,374
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/154,437
PRIOR FILING DATE: 1999-09-17
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 162
TYPE: PRT
ORGANISM: Piscirickettsia salmonis
US-10-261-446-2

Query Match 61.0%; Score 815; DB 28; Length 162;
Best Local Similarity 98.8%; Pred. No. 1.5e-76;
Matches 158; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 97 RGCLQGSSLLIIISVFLVGCNQFNRQEVGAATGAVVGGVAGQLFGKSGRVSMAIGGAVL 156
DB 3 RGCLQGSSLLIIISVFLVGCNQFNRQEVGAATGAVVGGVAGQLFGKSGRVSMAIGGAVL 62
QY 157 GGLGSKIGQSMDDQDKIKLQNSLEKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKOE 216
DB 63 GGLGSKIGQSMDDQDKIKLQNSLEKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKOE 122
QY 217 RQOYCREPQOKAMTAGQKOEIYGTACPODGRWQVISTEK 256
DB 123 RQOYCREPQOKAMTAGQKOEIYGTACPODGRWQVISTEK 162

RESULT 7
PCT-US00-00708-4
Sequence 4, Application PC/TUS00000708
GENERAL INFORMATION:
APPLICANT: Levy, Ilan
APPLICANT: Shoseyov, Oded
APPLICANT: Nussinovitch, Amos
TITLE OF INVENTION: MODIFICATION OF POLYSACCHARIDE CONTAINING MATERIALS
FILE REFERENCE: 00/20910
CURRENT APPLICATION NUMBER: PCT/US00/00708
CURRENT FILING DATE: 2000-11-02
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4
LENGTH: 190
TYPE: PRT
ORGANISM: Clostridium cellulovorans
PCT-US00-00708-4

Query Match 36.5%; Score 487; DB 1; Length 190;
Best Local Similarity 46.8%; Pred. No. 5.2e-42;
Matches 116; Conservative 12; Mismatches 30; Indels 90; Gaps 7;
QY 1 MSVEFYNSKSAQTNSITPIIKITWSDSLNLDVKRYVYTSDBGTQGTQCFWCDHAGAL 60

Db 1 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVVKVRYVYTS DGTQGTQOTFWCDHAGAL 60
QY 61 LGNSYVDNTSKVTANFVKETASPTSTYDYLDPDSHMRGCLQGSSLLIIISVFLVGCQAQNF 120
Db 61 LGNSYVDNTSKVTANFVKETASPTSTYDYLDPDSHMRGCLQGSSLLIIISVFLVGCQAQNF 92
QY 121 RQEVGAATGAVGVAGVAGQFPGKSGRVSMAIGAVLGLGSGKIGOSMDQDQKIKLNQSL 180
Db 93 RQEVGAATGAVGVAGVAGVAGQFPGKSGRVSMAIGAVLGLGSGKIGOSMDQDQKIKLNQSL 111
QY 181 EKVKGAVTR--WRNPDTGNSYSVEPVRTYQRYNKRQYCRFQKQKAMIAQKQEIY 238
Db 112 EKVKGAVTR--WRNPDTGNSYSVEPVRTYQRYNKRQYCRFQKQKAMIAQKQEIY 238
QY 239 GTACPOPD 246
Db 154 GTA-PGPD 160

RESULT 8
PCT-US00-00708-10
; Sequence 10, Application PC/TUS00000708
; GENERAL INFORMATION:
; APPLICANT: Levy, ilan
; APPLICANT: Shoseyov, Oded
; APPLICANT: Nussinovitch, Amos
; TITLE OF INVENTION: MODIFICATION OF POLYSACCHARIDE CONTAINING MATERIALS
; FILE REFERENCE: 00/20910
; CURRENT APPLICATION NUMBER: PCT/US00/00708
; CURRENT FILING DATE: 2000-11-02
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 326
; TYPE: PRT
; ORGANISM: recombinant protein sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (30)..(208)
; OTHER INFORMATION: taken from Clostridium cellulovorans
; NAME/KEY: misc feature
; LOCATION: (226)..(326)
; OTHER INFORMATION: taken from bovine
PCT-US00-00708-10

Query Match 36.5%; Score 487; DB 1; Length 326;
Best Local Similarity 46.8%; Pred. No. 1.1e-41;
Matches 116; Conservative 12; Mismatches 30; Indels 90; Gaps 7;
QY 1 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVVKVRYVYTS DGTQGTQOTFWCDHAGAL 60
Db 35 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVVKVRYVYTS DGTQGTQOTFWCDHAGAL 94
QY 61 LGNSYVDNTSKVTANFVKETASPTSTYDYLDPDSHMRGCLQGSSLLIIISVFLVGCQAQNF 120
Db 95 LGNSYVDNTSKVTANFVKETASPTSTYDYLDPDSHMRGCLQGSSLLIIISVFLVGCQAQNF 126
QY 121 RQEVGAATGAVGVAGVAGQFPGKSGRVSMAIGAVLGLGSGKIGOSMDQDQKIKLNQSL 180
Db 127 RQEVGAATGAVGVAGVAGVAGQFPGKSGRVSMAIGAVLGLGSGKIGOSMDQDQKIKLNQSL 145
QY 181 EKVKGAVTR--WRNPDTGNSYSVEPVRTYQRYNKRQYCRFQKQKAMIAQKQEIY 238
Db 146 EKVKGAVTR--WRNPDTGNSYSVEPVRTYQRYNKRQYCRFQKQKAMIAQKQEIY 238
QY 239 GTACPOPD 246
Db 188 GTA-PGPD 194

RESULT 9
PCT-US00-00708-6

; Sequence 6, Application PC/TUS00000708
; GENERAL INFORMATION:
; APPLICANT: Levy, ilan
; APPLICANT: Shoseyov, Oded
; APPLICANT: Nussinovitch, Amos
; TITLE OF INVENTION: MODIFICATION OF POLYSACCHARIDE CONTAINING MATERIALS
; FILE REFERENCE: 00/20910
; CURRENT APPLICATION NUMBER: PCT/US00/00708
; CURRENT FILING DATE: 2000-11-02
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Clostridium cellulovorans
PCT-US00-00708-6

Query Match 36.5%; Score 487; DB 1; Length 340;
Best Local Similarity 46.8%; Pred. No. 1.2e-41;
Matches 116; Conservative 12; Mismatches 30; Indels 90; Gaps 7;
QY 1 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVVKVRYVYTS DGTQGTQOTFWCDHAGAL 60
Db 1 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVVKVRYVYTS DGTQGTQOTFWCDHAGAL 60
QY 61 LGNSYVDNTSKVTANFVKETASPTSTYDYLDPDSHMRGCLQGSSLLIIISVFLVGCQAQNF 120
Db 61 LGNSYVDNTSKVTANFVKETASPTSTYDYLDPDSHMRGCLQGSSLLIIISVFLVGCQAQNF 92
QY 121 RQEVGAATGAVGVAGVAGQFPGKSGRVSMAIGAVLGLGSGKIGOSMDQDQKIKLNQSL 180
Db 93 RQEVGAATGAVGVAGVAGVAGQFPGKSGRVSMAIGAVLGLGSGKIGOSMDQDQKIKLNQSL 111
QY 181 EKVKGAVTR--WRNPDTGNSYSVEPVRTYQRYNKRQYCRFQKQKAMIAQKQEIY 238
Db 112 EKVKGAVTR--WRNPDTGNSYSVEPVRTYQRYNKRQYCRFQKQKAMIAQKQEIY 238
QY 239 GTACPOPD 246
Db 154 GTA-PGPD 160

RESULT 10
PCT-US00-00708-2
; Sequence 2, Application PC/TUS00000708
; GENERAL INFORMATION:
; APPLICANT: Levy, ilan
; APPLICANT: Shoseyov, Oded
; APPLICANT: Nussinovitch, Amos
; TITLE OF INVENTION: MODIFICATION OF POLYSACCHARIDE CONTAINING MATERIALS
; FILE REFERENCE: 00/20910
; CURRENT APPLICATION NUMBER: PCT/US00/00708
; CURRENT FILING DATE: 2000-11-02
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Clostridium cellulovorans
PCT-US00-00708-2

Query Match 36.2%; Score 483; DB 1; Length 163;
Best Local Similarity 51.7%; Pred. No. 1.1e-41;
Matches 106; Conservative 9; Mismatches 14; Indels 76; Gaps 4;
QY 1 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVVKVRYVYTS DGTQGTQOTFWCDHAGAL 60
Db 7 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVVKVRYVYTS DGTQGTQOTFWCDHAGAL 66
QY 61 LGNSYVDNTSKVTANFVKETASPTSTYDYLDPDSHMRGCLQGSSLLIIISVFLVGCQAQNF 120
Db 67 LGNSYVDNTSKVTANFVKETASPTSTYDYLDPDSHMRGCLQGSSLLIIISVFLVGCQAQNF 98

```

QY 121 ROEUGAATGAVGVAGOLFPGKSGRVSMIAIGGAVLGLIGSKIGQSMDDQDKIKLNQSL 180
Db 99 -----FGFASGRATL-----KKQGFITIQ----- 117
QY 181 EKVAGQVTR--WRNPDTGNSVSVE 203
Db 118 -----GRITKSDWSNYTQTDNSFD 137

RESULT 11
PCT-US00-00708-8
; Sequence 8, Application PC/TUS00000708
; GENERAL INFORMATION:
; APPLICANT: Levy, Ilan
; APPLICANT: Shoseyov, Oded
; APPLICANT: Nussinovitch, Amos
; TITLE OF INVENTION: MODIFICATION OF POLYSACCHARIDE CONTAINING MATERIALS
; FILE REFERENCE: 00/20910
; CURRENT APPLICATION NUMBER: PCT/US00/00708
; CURRENT FILING DATE: 2000-11-02
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 426
; TYPE: PRT
; ORGANISM: recombinant protein sequence
; NAME/KEY: misc feature
; LOCATION: (1)..(263)
; OTHER INFORMATION: protein A from cloning vector
; NAME/KEY: misc feature
; LOCATION: (265)..(426)
; OTHER INFORMATION: CBPA
PCT-US00-00708-8

Query Match 36.2%; Score 483; DB 1; Length 426;
Best Local Similarity 51.7%; Pred. No. 4.4e-41;
Matches 106; Conservative 9; Mismatches 14; Indels 76; Gaps 4;

QY 1 MSVEFYNSKSAQNTSIPIIKITNTSDSLNLDNDVKRYVYVTSYDGTQGTQFVCDHAGAL 60
Db 270 MSVEFYNSKSAQNTSIPIIKITNTSDSLNLDNDVKRYVYVTSYDGTQGTQFVCDHAGAL 329
QY 61 LGNSYVDNTSKVTANFVKETASPTSTYDYLDPHSHMRGCLGSSLIISVFLVGCAQNF 120
Db 330 LGNSYVDNTSKVTANFVKETASPTSTYDYLDPHSHMRGCLGSSLIISVFLVGCAQNF 361
QY 121 ROEUGAATGAVGVAGOLFPGKSGRVSMIAIGGAVLGLIGSKIGQSMDDQDKIKLNQSL 180
Db 362 -----FGFASGRATL-----KKQGFITIQ----- 380
QY 181 EKVAGQVTR--WRNPDTGNSVSVE 203
Db 381 -----GRITKSDWSNYTQTDNSFD 400

RESULT 12
US-09-808-898-22
; Sequence 22, Application US/09808898
; GENERAL INFORMATION:
; APPLICANT: Bryan, Bruce
; APPLICANT: Szent-Gyorgyi, Christopher
; APPLICANT: Szecpaniak, William
; TITLE OF INVENTION: RENILIA RENIFORMIS FLOURESCENT PROTEINS, NUCLEIC ACIDS ENCODING T
; TITLE OF INVENTION: FLOURESCENT PROTEINS AND THE USE THEREOF IN DIAGNOSTICS, HIGH TH
; TITLE OF INVENTION: SCREENING AND NOVELTY ITEMS
; FILE REFERENCE: 24729-128
; CURRENT APPLICATION NUMBER: US/09/808,898
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 60/189,691
; PRIOR FILING DATE: 2000-03-15
; PRIOR APPLICATION NUMBER: 09/277,716
; PRIOR FILING DATE: 1999-03-26

```

```

; PRIOR APPLICATION NUMBER: 08/757,046
; PRIOR FILING DATE: 1996-11-25
; PRIOR APPLICATION NUMBER: 08/597,274
; PRIOR FILING DATE: 1996-02-06
; PRIOR APPLICATION NUMBER: 08/908,909
; PRIOR FILING DATE: 1997-08-08
; PRIOR APPLICATION NUMBER: 08/990,103
; PRIOR FILING DATE: 1997-12-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Caussia
US-09-808-898-22

Query Match 36.0%; Score 480.5; DB 23; Length 382;
Best Local Similarity 62.6%; Pred. No. 7e-41;
Matches 107; Conservative 10; Mismatches 33; Indels 21; Gaps 4;

QY 1 MSVEFYNSKSAQNTSIPIIKITNTSDSLNLDNDVKRYVYVTSYDGTQGTQFVCDHAGAL 60
Db 1 MSVEFYNSKSAQNTSIPIIKITNTSDSLNLDNDVKRYVYVTSYDGTQGTQFVCDHAGAL 60
QY 61 LGNSYVDNTSKVTANFVKETASPTSTYDYLDPHSHMRGCLGSSLIISVFLVGCAQ 117
Db 61 LGNSYVDNTSKVTANFVKETASPTSTYDYLDPHSHMRGCLGSSLIISVFLVGCAQ 120
QY 118 NFSRQ-----EVGAATGAVGVAGOLFPGKSGRVSMIAIGGAVLGLIGS 162
Db 121 NYTQTDNSYDFDASSSTPVVNPVKTGY-----IGGAKVLGTAPGS 159

RESULT 13
US-08-330-394-2
; Sequence 2, Application US/08330394
; GENERAL INFORMATION:
; APPLICANT: Shoseyov, Oded
; APPLICANT: Yosef, Karmey
; APPLICANT: Shpiegl, Itai
; APPLICANT: Goldstein, Marc A.
; APPLICANT: Doi, Roy H.
; TITLE OF INVENTION: METHODS OF DETECTION USING THE
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 AVENUE OF THE AMERICAS
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/330,394
; FILING DATE: 27-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CIP OF PCT/US94/04132
; FILING DATE: 14-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REFERENCE/DOCKET NUMBER: 7809-005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864
; TELEX: 66441 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 162 amino acids

```

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-330-394-2

Query Match 35.7%; Score 477; DB 7; Length 162;
Best Local Similarity 97.8%; Pred. No. 4.7e-41;
Matches 90; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVVKVRYYYTSDGTQGTQGTWCDHAGAL 60
Db 6 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVVKVRYYYTSDGTQGTQGTWCDHAGAL 65
QY 61 LGNSYVDNTSKVTANFVKETASPTSTYDYILD 92
Db 66 LGNSYVDNTSKVTANFVKETASPTSTYDYIVE 97

RESULT 14

US-08-962-588-2
; Sequence 2, Application US/08962588
; GENERAL INFORMATION:
; APPLICANT: Shoseyov, Oded
; APPLICANT: Spiegler, Etai
; APPLICANT: Roiz, Levava
; TITLE OF INVENTION: METHODS FOR USE OF A CELLULOSE
; TITLE OF INVENTION: BINDING DOMAIN
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 AVENUE OF THE AMERICAS
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: U.S.A.
; ZIP: 10036

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/962,588
; FILING DATE: To be assigned
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: 08/537,893
; FILING DATE:
; APPLICATION NUMBER: PCT/US94/04132
; FILING DATE: 14-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REFERENCE/DOCKET NUMBER: 7809-017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864
; TELEX: 66441 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 162 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-962-588-2

Query Match 35.7%; Score 477; DB 13; Length 162;
Best Local Similarity 97.8%; Pred. No. 4.7e-41;
Matches 90; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVVKVRYYYTSDGTQGTQGTWCDHAGAL 60
Db 6 MSVEFYNSKSAQTNSITPIIKITNTSDSLNLDVVKVRYYYTSDGTQGTQGTWCDHAGAL 65
QY 61 LGNSYVDNTSKVTANFVKETASPTSTYDYILD 92

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 20:13:32 ; Search time 14.5907 Seconds
(without alignments)
824.821 Million cell updates/sec

Title: US-09-677-374-6
Perfect score: 1335
Sequence: 1 MSVEFYNNKSAQNTSITPI.....IYGTAQPDGRWQVISTEK 256

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 248976 seqs, 47010500 residues

Total number of hits satisfying chosen parameters: 248976

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents AA New:*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1335	100.0	256	US-10-261-445B-6	Sequence 6, Appli
2	832	62.3	161	US-10-261-445B-4	Sequence 4, Appli
3	815	61.0	162	US-10-241-602B-6	Sequence 6, Appli
4	815	61.0	162	US-10-261-445B-2	Sequence 2, Appli
5	245	18.4	599	US-09-955-555B-29	Sequence 29, Appli
6	124	9.3	499	1 PCT-US03-19153-328	Sequence 328, App
7	112	8.4	20	US-10-261-445B-16	Sequence 16, Appl
8	102.5	7.7	223	US-09-581-286A-434	Sequence 434, App
9	102.5	7.7	230	US-09-581-286A-309	Sequence 309, App
10	94.5	7.1	106	US-09-897-516A-6133	Sequence 5133, Ap
11	87	6.5	363	US-10-425-114A-61312	Sequence 61312, A
12	86.5	6.5	99	US-10-389-647-605	Sequence 605, App
13	86.5	6.5	403	US-09-897-516A-7766	Sequence 7766, Ap
14	86.5	6.5	1117	1 PCT-US03-21510-114	Sequence 114, App
15	84	6.3	423	1 PCT-US02-39429-8	Sequence 8, Appli
16	83.5	6.3	683	US-60-487-610-1870	Sequence 1870, App
17	83.5	6.3	1397	1 PCT-US02-35624-174	Sequence 174, App
18	83	6.2	272	US-10-679-063-14893	Sequence 14893, A
19	83	6.2	296	US-10-425-114A-51273	Sequence 51273, A
20	83	6.2	345	US-10-425-114A-59685	Sequence 59685, A
21	83	6.2	521	US-10-425-114A-70443	Sequence 70443, A
22	83	6.2	3011	1 PCT-US03-19834-2	Sequence 2, Appli
23	82.5	6.2	497	US-10-679-063-15782	Sequence 15782, A
24	82.5	6.2	734	US-10-425-114A-49440	Sequence 49440, A
25	82.5	6.2	1327	1 PCT-US03-28626-27	Sequence 27, Appl
26	82.5	6.2	1327	US-10-661-398-27	Sequence 27, Appl

27	82.5	6.2	2599	5	US-09-674-546A-1669	Sequence 1669, Ap
28	82	6.1	595	6	US-10-679-063-23425	Sequence 23425, A
29	80.5	6.0	1261	6	US-10-679-063-16542	Sequence 16542, A
30	80	6.0	293	1	PCT-US03-28227-2863	Sequence 2863, Ap
31	80	6.0	515	7	US-60-478-196-3014	Sequence 3014, Ap
32	80	6.0	1016	1	PCT-US02-35624-95	Sequence 95, Appl
33	80	6.0	3033	6	US-10-009-002-5	Sequence 5, Appli
34	79.5	6.0	330	6	US-10-089-278-8	Sequence 8, Appli
35	79	5.9	435	6	US-10-425-114A-73075	Sequence 73075, A
36	79	5.9	593	6	US-10-435-696-50	Sequence 50, Appl
37	78.5	5.9	261	6	US-10-689-006-24	Sequence 24, Appl
38	78	5.8	222	5	US-09-408-020-58	Sequence 58, Appl
39	78	5.8	438	1	PCT-US02-39429-88	Sequence 88, Appl
40	78	5.8	515	6	US-10-679-063-22745	Sequence 22745, A
41	78	5.8	517	6	US-10-679-063-22744	Sequence 22744, A
42	78	5.8	852	1	PCT-US03-19153-120	Sequence 120, App
43	78	5.8	3011	6	US-10-296-734-406	Sequence 406, App
44	77.5	5.8	2871	7	US-60-502-656-256	Sequence 256, App
45	77	5.8	1755	1	PCT-US03-26780-344	Sequence 344, Ap

ALIGNMENTS

RESULT 1

US-10-261-445B-6
; Sequence 6, Application US/10261445B
; GENERAL INFORMATION:
; APPLICANT: Kuziyk, Michael A.
; APPLICANT: Burian, Jan
; APPLICANT: Kay, William W.
; APPLICANT: Thornton, Julian C.
; TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY IN FISH AGAINST
; TITLE OF INVENTION: RICKETTSIAL DISEASES, AND ASSOCIATED PREVENTATIVE THERAPY
; FILE REFERENCE: 4616-64095
; CURRENT APPLICATION NUMBER: US/10/261,445B
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US 09/677,374
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/154,437
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Piscirickettsia salmonis
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (-95)..(-1)
US-10-261-445B-6

Query Match 100.0%; Score 1335; DB 6; Length 256;
Best Local Similarity 100.0%; Pred. No. 1.5e-112;
Matches 256; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MSVEFYNNKSAQNTSITPIIKITNTSDNLNDVKRYVYTSDTGQTQCFWCDHAGAL 60
DB	1	MSVEFYNNKSAQNTSITPIIKITNTSDNLNDVKRYVYTSDTGQTQCFWCDHAGAL 60
QY	61	LGNSYVDNTSKVTANFVKETASPTSTYDTYLDPSHMRGCLQGSSLIISVFLVGCAQNF 120
DB	61	LGNSYVDNTSKVTANFVKETASPTSTYDTYLDPSHMRGCLQGSSLIISVFLVGCAQNF 120
QY	121	RQEVGAATGAVVGGVAGQLFGKSGRVSNAIGAVLGLIGSKIGQSDMDQDKIKLNQSL 180
DB	121	RQEVGAATGAVVGGVAGQLFGKSGRVSNAIGAVLGLIGSKIGQSDMDQDKIKLNQSL 180
QY	181	EKVKGAGQTRWRNPDTGNSYVEPVRTYQRYNKQRRQYCFRFOQKAMIAQKQEIYCT 240
DB	181	EKVKGAGQTRWRNPDTGNSYVEPVRTYQRYNKQRRQYCFRFOQKAMIAQKQEIYCT 240
QY	241	ACQPQDGRWQVISTEK 256

```
Db      241 ACPQPDGRWQVISTEK 256
|||||
Query Match      62.3%; Score 832; DB 6; Length 161;
Best Local Similarity 100.0%; Pred. No. 1.5e-67;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      96 MRGCLQGSSLIISVFLVGCACQNFSEQVGAATGAVVGGVAGOLFQKSGSRVSMAGGAV 155
Db      1 MRGCLQGSSLIISVFLVGCACQNFSEQVGAATGAVVGGVAGOLFQKSGSRVSMAGGAV 60
QY      156 LGLGSKIGQSMDDQDKIKLQSLKVKAGQVTRWRNPDGTGNSYSVEPVRTYQRYNKOE 215
Db      61 LGLGSKIGQSMDDQDKIKLQSLKVKAGQVTRWRNPDGTGNSYSVEPVRTYQRYNKOE 120
QY      216 RQCYCREPQQKAMIAQKQEIYGTACPOPDGRWQVISTEK 256
Db      121 RQCYCREPQQKAMIAQKQEIYGTACPOPDGRWQVISTEK 161

RESULT 3
US-10-241-602B-6
; Sequence 6, Application US/10241602B
; GENERAL INFORMATION:
; APPLICANT: Simard, Nathalie
; APPLICANT: Brouwers, Huub
; APPLICANT: Jones, Simon
; APPLICANT: Griffiths, Steve
; APPLICANT: Valenzuela, Pablo
; APPLICANT: Burzio, Luis
; TITLE OF INVENTION: Sequences from Piscirickettsia salmonis
; FILE REFERENCE: H-32319A
; CURRENT APPLICATION NUMBER: US/10/241,602B
; CURRENT FILING DATE: 2002-09-11
; PRIOR APPLICATION NUMBER: PCT/GB01/01055
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: GB0005838.8
; PRIOR FILING DATE: 2000-03-11
; PRIOR APPLICATION NUMBER: GB0016080.4
; PRIOR FILING DATE: 2000-07-01
; PRIOR APPLICATION NUMBER: GB0016082.0
; PRIOR FILING DATE: 2000-07-01
; PRIOR APPLICATION NUMBER: GB0018599.1
; PRIOR FILING DATE: 2000-07-29
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6

Db      241 ACPQPDGRWQVISTEK 256
|||||
Query Match      61.0%; Score 815; DB 6; Length 162;
Best Local Similarity 98.8%; Pred. No. 5.1e-66;
Matches 158; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      97 RGCLOGSSLIISVFLVGCACQNFSEQVGAATGAVVGGVAGOLFQKSGSRVSMAGGAVL 156
Db      3 RGCLOGSSLIISVFLVGCACQNFSEQVGAATGAVVGGVAGOLFQKSGSRVSMAGGAVL 62
QY      157 GGLGSKIGQSMDDQDKIKLQSLKVKAGQVTRWRNPDGTGNSYSVEPVRTYQRYNKOER 216
Db      63 GGLGSKIGQSMDDQDKIKLQSLKVKAGQVTRWRNPDGTGNSYSVEPVRTYQRYNKOER 122
QY      217 RQCYCREPQQKAMIAQKQEIYGTACPOPDGRWQVISTEK 256
Db      123 RQCYCREPQQKAMIAQKQEIYGTACPOPDGRWQVISTEK 162

RESULT 4
US-10-261-445B-2
; Sequence 2, Application US/10261445B
; GENERAL INFORMATION:
; APPLICANT: Kuzyk, Michael A.
; APPLICANT: Burian, Jan
; APPLICANT: Kay, William W.
; APPLICANT: Thornton, Julian C.
; TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY IN FISH AGAINST
; TITLE OF INVENTION: RICKETTSIAL DISEASES, AND ASSOCIATED PREVENTATIVE THERAPY
; FILE REFERENCE: 4616-64095
; CURRENT APPLICATION NUMBER: US/10/261,445B
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US 09/677,374
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/154,437
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Piscirickettsia salmonis
US-10-261-445B-4

Query Match      62.3%; Score 832; DB 6; Length 161;
Best Local Similarity 100.0%; Pred. No. 1.5e-67;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      96 MRGCLQGSSLIISVFLVGCACQNFSEQVGAATGAVVGGVAGOLFQKSGSRVSMAGGAV 155
Db      1 MRGCLQGSSLIISVFLVGCACQNFSEQVGAATGAVVGGVAGOLFQKSGSRVSMAGGAV 60
QY      156 LGLGSKIGQSMDDQDKIKLQSLKVKAGQVTRWRNPDGTGNSYSVEPVRTYQRYNKOE 215
Db      61 LGLGSKIGQSMDDQDKIKLQSLKVKAGQVTRWRNPDGTGNSYSVEPVRTYQRYNKOE 120
QY      216 RQCYCREPQQKAMIAQKQEIYGTACPOPDGRWQVISTEK 256
Db      121 RQCYCREPQQKAMIAQKQEIYGTACPOPDGRWQVISTEK 161

RESULT 5
US-09-955-555B-29
; Sequence 29, Application US/09955555B
; GENERAL INFORMATION:
; APPLICANT: Bott, Richard R.
; APPLICANT: Clarkson, Kathleen A.
; APPLICANT: Fowler, Timothy
; APPLICANT: Liu, Chung-Cheng
; APPLICANT: Ward, Michael
```

```

; APPLICANT: Xia, Hai-Ying
; TITLE OF INVENTION: Enzymatic Array and Process of Making
; FILE REFERENCE: Same
; CURRENT APPLICATION NUMBER: US/09/955,555B
; CURRENT FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: US 08/559,958
; PRIOR FILING DATE: 1995-11-17
; PRIOR APPLICATION NUMBER: US 60/005,701
; PRIOR FILING DATE: 1995-10-17
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29
; LENGTH: 599
; TYPE: PRT
; ORGANISM: Clostridium thermocellum
US-09-955-555B-29

```

```

Query Match      18.4%; Score 245; DB 5; Length 599;
Best Local Similarity 51.5%; Pred. No. 5.4e-14;
Matches 53; Conservative 16; Mismatches 30; Indels 4; Gaps 3;

QY 1 MSVEFVNSKSAQTNSITPIIKITNTSDLDNLNDVKRYYYTSDGTQGTFCWDHAGAL 60
DB 384 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
      1 MSVEFVNSKSAQTNSITPIIKITNTSDLDNLNDVKRYYYTSDGTQGTFCWDHAGAL 60
      384 LKVEFYNSPSTTNSINPQFKVTNTGSSAIDLSKLTLLRYYYTVDGQKQDTCWDHA-AI 442

QY 61 LGN--SYVDNTSKVTANFVKETASPTSTYDTYLDPSHMRGCLQ 101
DB 443 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
      61 LGN--SYVDNTSKVTANFVKETASPTSTYDTYLDPSHMRGCLQ 101
      443 IGSNGSYNGITSNVKGTFFVMSSS-TNNADTYLISFTGTTLE 484

```

```

RESULT 6
PCT-US03-19153-328
; Sequence 328, Application PC/TUS0319153
; GENERAL INFORMATION:
; APPLICANT: Diversa Corporation
; TITLE OF INVENTION: XYLANASES, NUCLEIC ACIDS ENCODING THEM AND
; FILE REFERENCE: 09010-290W01
; CURRENT APPLICATION NUMBER: PCT/US03/19153
; CURRENT FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/389,299
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 378
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 328
; LENGTH: 499
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Obtained from an environmental sample.
; NAME/KEY: SIGNAL
; LOCATION: (1)...(29)
PCT-US03-19153-328

```

```

Query Match      9.3%; Score 124; DB 1; Length 499;
Best Local Similarity 33.0%; Pred. No. 0.0033;
Matches 33; Conservative 14; Mismatches 43; Indels 10; Gaps 3;

QY 1 MSVEFVNSKSAQTNSITPIIKITNTSDLDNLNDVKRYYYTSDGTQGTFCWDHAGAL 60
DB 354 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
      1 MSVEFVNSKSAQTNSITPIIKITNTSDLDNLNDVKRYYYTSDGTQGTFCWDHAGAL 60
      354 LSVQRTGDSVNSQIRPQIHVKNSKTTVNLKNVTYRYWNTK-NKGQNFCDYAKI- 411

QY 61 LGNSVVDNTSKVTANFVKETASPTSTYDTYLDPSHMRGCL 100
DB 412 :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
      61 LGNSVVDNTSKVTANFVKETASPTSTYDTYLDPSHMRGCL 100
      412 -----GCSNVTHRFV-TLQKPVKGADAYLELGFNGTLE 443

```

```

RESULT 7
US-10-261-445B-16
; Sequence 16, Application US/10261445B
; GENERAL INFORMATION:

```

```

; APPLICANT: Kuzyk, Michael A.
; APPLICANT: Burian, Jan
; APPLICANT: Kay, William W.
; APPLICANT: Thornton, Julian C.
; TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY IN FISH AGAINST
; FILE REFERENCE: 4616-64095
; CURRENT APPLICATION NUMBER: US/10/261,445B
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US 09/677,374
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/154,437
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Piscirickettsia salmonis
US-10-261-445B-16

```

```

Query Match      8.4%; Score 112; DB 6; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.0006;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 204 PVRTYQRYNKQERRQQCYCRE 223
DB 1 :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
      204 PVRTYQRYNKQERRQQCYCRE 223
      1 PVRTYQRYNKQERRQQCYCRE 20

```

```

RESULT 8
US-09-581-286A-434
; Sequence 434, Application US/09581286A
; GENERAL INFORMATION:
; APPLICANT: ROSS, BRUCE C.
; APPLICANT: BARR, IAN G.
; APPLICANT: PATTERSON, MICHELLE A.
; APPLICANT: AGIUS, CATHERINE T.
; APPLICANT: ROTH, LINDA J.
; APPLICANT: MARGETTS, MAL B.
; APPLICANT: HOCKING, DIANNA M.
; APPLICANT: WEBB, ELIZABETH A.
; TITLE OF INVENTION: PORPHYROMONAS GINGIVALIS POLYPEPTIDES AND NUCLEOTIDES
; FILE REFERENCE: 4137-3
; CURRENT APPLICATION NUMBER: US/09/581,286A
; CURRENT FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: PCT/AU98/01023
; PRIOR FILING DATE: 1998-12-10
; PRIOR APPLICATION NUMBER: AU PP 0839
; PRIOR FILING DATE: 1997-12-10
; PRIOR APPLICATION NUMBER: AU PP 1182
; PRIOR FILING DATE: 1997-12-31
; PRIOR APPLICATION NUMBER: AU PP 1846
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: AU PP 2264
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: AU PP 2911
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: AU PP 3128
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: AU PP 3338
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: AU PP 3654
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: AU PP 4917
; PRIOR FILING DATE: 1998-07-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 721
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 434
; LENGTH: 223
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis

```


QY 30 DLNLNDVKVRYYYTSDGTQGTQGTFCWDHAGALLGNSYVDNTSKVTANFVKETASPTSTYDT 89

; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 690022.54701PC
; CURRENT APPLICATION NUMBER: PCT/US02/39429
; CURRENT FILING DATE: 2002-12-06
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 423
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Predicted amino acid sequence of B9E9 single chain antibody- genc
PCT/US02/39429-8

Query Match	6.3%;	Score 84;	DB 1;	Length 423;
Best Local Similarity	25.2%;	Pred. No. 10;		
Matches	55;	Conservative	22;	Mismatches 73; Indels 68; Gaps 10;

QY	92	DPSHMRGCLQSSLLIIISVFLVCAQ-----NFSROEV-GAATGAVVGGVAGQLFGKSGG	145
DB	75	DKSSSTAYNQLSLTSSEDSAVVYCARAQLRPNNWYFDVWGAGTTVTVSKISG--GGSGG	131
QY	146	RVSMAIGGAVLGGIGSKIGSMDQDDKIKLNQSL-----EKV-----	183
DB	132	GGSGGGSGGGGGGGG--GGSSD-----IVLSQSPAILSPGKVTMTTCRASSSVSYMH	184
QY	184	-----KAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQERRQOYCRE	223
DB	185	WYQKPGSSPKWIVATSNLASGVPARFSGSGSGTSYSL-----TISRVEAEDAATYYCQ	240
QY	224	F--QQKAMLAGOKQEI-----YGTACFPQDGRWQVISTE	255
DB	241	WISNPPTFGAGTKLELKSSGSGSADPSKDSKAQVSAAE	278

Search completed: November 5, 2003, 20:25:59
Job time : 15.5907 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 20:06:21 ; Search time 29.3782 Seconds
(without alignments)
875.264 Million cell updates/sec

Title: US-09-677-374-2
Perfect score: 836
Sequence: 1 MNRGCLQSSLIISVFLVG.....IYGTAQRQDGRWQVISTEK 162

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues
Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A.Geneseq 15Jun03:*

1:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	836	100.0	162	22	AA1980
2	836	100.0	162	22	AA1981
3	836	100.0	162	23	AA1982
4	815	97.5	161	22	AA1983
5	815	97.5	161	23	AA1984
6	815	97.5	256	22	AA1985
7	815	97.5	256	23	AA1986
8	128.5	15.4	182	24	AA1987
9	112	13.4	20	22	AA1988

10	112	13.4	20	23	AAU97871	Outer surface lipo
11	108	12.9	197	23	AAO17565	M catarrhalis MCA1
12	108	12.9	224	22	AAO20105	Moraxella catarrha
13	108	12.9	224	23	ABG80432	Moraxella catarrha
14	105.5	12.6	223	20	ABY34487	Porphyromonas ging
15	105.5	12.6	230	20	ABY34362	Porphyromonas ging
16	102	12.2	154	11	ABR05799	PBOMP-2 gene prod.
17	102	12.2	155	23	ABG80423	Haemophilus influe
18	98	11.7	309	22	ABG15906	Novel human diagno
19	91	10.9	528	22	ABG26111	Spider recombinant
20	86.5	10.3	2309	22	ABG66232	Drosophila melanog
21	83.5	10.0	71	24	ABP58756	Rickettsia rickett
22	82	9.8	666	22	ABP58019	Drosophila melanog
23	81.5	9.7	651	20	AAU40097	Spider silk protei
24	81.5	9.7	651	23	AAU11781	Spider natural sil
25	81.5	9.7	718	12	AAU14308	N.clavipes draglin
26	81.5	9.7	718	19	AAWS3346	Nephila clavipes s
27	81.5	9.7	718	21	AAU59070	N. clavipes spider
28	80.5	9.6	102	22	ABP37981	Human polypeptide
29	80.5	9.6	170	22	ABP37981	Human GS930284 pro
30	80.5	9.6	302	22	AAU40157	Human polypeptide
31	80.5	9.6	302	23	AAU74394	Human ovarian tumo
32	80.5	9.6	318	22	ABP37982	Human GS930284 pro
33	80.5	9.6	542	22	ABP65790	Drosophila melanog
34	80.5	9.6	542	22	ABP65791	Drosophila melanog
35	80.5	9.6	542	22	ABP70501	Drosophila melanog
36	80.5	9.6	1640	23	ABP54727	Lactococcus lactis
37	80	9.6	116	19	AAU11028	H. pylori ORF 01cp
38	80	9.6	240	22	ABP0680	N. gonorrhoeae ami
39	80	9.6	2017	22	ABG06301	Novel human diagno
40	80	9.6	2599	21	AAU75098	Neisseria meningit
41	79.5	9.5	553	23	ABG79662	Invertebrate forag
42	79.5	9.5	553	24	ABR47395	Breast cancer asso
43	79.5	9.5	553	24	ABU07492	Protein differenti
44	79	9.4	112	23	ABP28121	Streptococcus poly
45	79	9.4	618	21	AAU56803	Human prostate can

ALIGNMENTS

RESULT 1
AAG78025
ID AAG78025 standard; Protein; 162 AA.
XX
AC AAG78025;
XX
XX
DT 15-JAN-2002 (first entry)
XX
DE Piscirickettsia salmonis polypeptide P10.6.
XX
KW Piscirickettsia salmonis; piscirickettsiosis; salmonid rickettsial;
KW septicemia; SRS; surface antigen; vaccine; antibacterial; fish;
KW ATCC VR-1361.
XX
OS Piscirickettsia salmonis.
XX
PN WO200158865-A2.
XX
PD 20-SEP-2001.
XX
PF 12-MAR-2001; 2001WO-GB01055.
XX
PR 11-MAR-2000; 2000GB-0005838.
PR 01-JUL-2000; 2000GB-0016080.
PR 01-JUL-2000; 2000GB-0016082.
XX
PR 29-JUL-2000; 2000GB-0018599.
XX
PA (AQUA-) AQUA HEALTH EURO LTD.
XX
PI Simard N, Brouwers H, Jones S, Griffiths S, Valenzuela P;
PI Burzio L;

```

DR WPI: 2001-639050/73.
DR N-PSDB; AAH79040.
XX
PT New nucleic acids encoding an amino acid sequence homologous to the
PT surface antigen present on Piscirickettsia salmonis are useful to
PT protect fish against piscirickettsiosis.
XX
XX Claim 6; Fig 5; 25pp; English.
XX
XX The invention relates to nucleic acid sequences and the encoded protein
CC of a least part of the surface antigen present on Piscirickettsia
CC salmonis for production of a vaccine with antibacterial activity to
CC protect fish against P. salmonis which causes piscirickettsiosis, also
CC known as salmonid rickettsial septicaemia.
XX
XX SQ Sequence 162 AA;
Query Match 100.0%; Score 836; DB 22; Length 162;
Best Local Similarity 100.0%; Pred. No. 3.8e-80;
Matches 162; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MNRGCLQGSSLLIIISVFLVGCANFQSRQEVGAATGAVVGGVAGQLFGKSGRVAIAIGGA 60
Db 1 MNRGCLQGSSLLIIISVFLVGCANFQSRQEVGAATGAVVGGVAGQLFGKSGRVAIAIGGA 60
QY 61 VLGGILGSKIGQSMDDQDKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQ 120
Db 61 VLGGILGSKIGQSMDDQDKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQ 120
QY 121 ERQQYCREFOQKAMIAGQKQEIYGTACRQPDGRWQVISTEK 162
Db 121 ERQQYCREFOQKAMIAGQKQEIYGTACRQPDGRWQVISTEK 162

RESULT 2
AAB81126
ID AAB81126 standard; Protein; 162 AA.
XX
AC AAB81126;
XX
DT 11-JUL-2001 (first entry)
XX
DE OSPA antigen amino acid sequence.
XX
KW Poikilothermic fish; Piscirickettsia salmonis; rickettsial pathogen;
KW vaccine; OSPA; salmonid rickettsial septicaemia; rickettsial disease;
KW SRS.
XX
OS Piscirickettsia salmonis.
XX
PH Key Location/Qualifiers
FT Region 110..129
FT /label= B_cell_epitope
XX
XX CA2281913-A1.
XX
XX 17-MAR-2001.
XX
XX 17-SEP-1999; 99CA-2281913.
XX
XX 17-SEP-1999; 99CA-2281913.
XX
XX (KAYW/) KAY W W.
XX
XX (BURI/) BURIAN J.
XX
XX (KUZY/) KUZYSK M A.
XX
XX Kay WW, Burian J, Kuzyk MA;
XX
XX WPI: 2001-316844/34.
XX
XX N-PSDB; AAF86246.
XX
XX Method for protecting poikilothermic fish against salmonid rickettsial
XX septicaemia and other rickettsial diseases comprises administering a

```

```

PT vaccine containing the OSPA protein of Piscirickettsia salmonis -
XX
XX Example 2; Fig 2B; 35pp; English.
XX
XX This invention relates to a method for the protection against infection
CC of a poikilothermic fish by the bacterial pathogen, Piscirickettsia
CC salmonis. The method comprises administering an immunogenic amount of a
CC P. salmonis specific antigen termed OSPA, or an immunogenic fragment of
CC OSPA in the form of a vaccine. The method is used for protecting animals,
CC particularly poikilothermic fish, against the bacterial pathogen
CC P. salmonis. The method is also useful for protecting against salmonid
CC rickettsial septicaemia (SRS) and other rickettsial diseases. The present
CC sequence represents P. salmonis OSPA protein. An OSPA protein with an
CC N-terminal fusion partner is used in a vaccine to create an anti-OSPA
XX antibody response.
XX
XX SQ Sequence 162 AA;
Query Match 100.0%; Score 836; DB 22; Length 162;
Best Local Similarity 100.0%; Pred. No. 3.8e-80;
Matches 162; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MNRGCLQGSSLLIIISVFLVGCANFQSRQEVGAATGAVVGGVAGQLFGKSGRVAIAIGGA 60
Db 1 MNRGCLQGSSLLIIISVFLVGCANFQSRQEVGAATGAVVGGVAGQLFGKSGRVAIAIGGA 60
QY 61 VLGGILGSKIGQSMDDQDKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQ 120
Db 61 VLGGILGSKIGQSMDDQDKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQ 120
QY 121 ERQQYCREFOQKAMIAGQKQEIYGTACRQPDGRWQVISTEK 162
Db 121 ERQQYCREFOQKAMIAGQKQEIYGTACRQPDGRWQVISTEK 162

RESULT 3
AAU97867
ID AAU97867 standard; Protein; 162 AA.
XX
AC AAU97867;
XX
DT 12-AUG-2002 (first entry)
XX
DE Piscirickettsia salmonis outer surface lipoprotein OSPA.
XX
KW Outer surface lipoprotein; OSPA; antibacterial; immunosuppressive;
KW vaccine; poikilothermic fish; fin-fish; Rickettsial septicaemia;
KW Rickettsial disease.
XX
OS Piscirickettsia salmonis.
XX
XX CA2339327-A1.
XX
XX 15-MAR-2002.
XX
XX 19-MAR-2001; 2001CA-2339327.
XX
XX 15-SEP-2000; 2000US-0677374.
XX
XX (THOR/) THORNTON J C.
XX
XX (KAYW/) KAY W W.
XX
XX (BURI/) BURIAN J.
XX
XX (KUZY/) KUZYSK M A.
XX
XX Thornton JC, Kay WW, Burian J, Kuzyk MA;
XX
XX WPI: 2002-455221/49.
XX
XX N-PSDB; ABK52401.
XX
XX Inducing immunity in fin fish to Rickettsial septicaemia, comprises
PT administration of an outer surface lipoprotein (OSPA) of a bacterial
PT strain, as a vaccine -
XX

```

PS Claim 15; Fig 2; 55pp; English.

XX The invention describes a method of protecting a poikilothermic fish
CC against infection by the bacterial pathogen *Piscirickettsia salmonis*
CC comprising administering either intraperitoneally, by immersion or
CC orally, an immunogenic amount of principal antigen, the *OspA* (outer
CC surface lipoprotein), its variants, non-lipidated form or antigenic
CC peptide derived or synthesized with or without an adjuvant. The new
CC method is used to provide an outer surface lipoprotein (*OspA*) of
CC bacterial strain *Piscirickettsia salmonis* as a vaccine to induce immunity
CC in fin-fish against *Rickettsial* septicemia and other related
CC *Rickettsial* diseases caused by either a virus, bacteria or parasite.
CC This is the amino acid sequence of the *Piscirickettsia salmonis* outer
CC surface lipoprotein, *OspA*, used in the creation of the vaccine described
CC in the invention.

XX Sequence 162 AA;

Query Match 100.0%; Score 836; DB 23; Length 162;
Best Local Similarity 100.0%; Pred. No. 3.8e-80;
Matches 162; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MNRGCLQGSSLLIIISVFLVGCQNFSRQEVGAATGAVVGVAGOLFSGSRVMAIGA 60
DB 1 MNRGCLQGSSLLIIISVFLVGCQNFSRQEVGAATGAVVGVAGOLFSGSRVMAIGA 60
QY 61 VLGGLIGSKIGSMDOODKIKLNQSLKVKAGQVTRWRNPDTGNSVSEVPRTYQYNKQ 120
DB 61 VLGGLIGSKIGSMDOODKIKLNQSLKVKAGQVTRWRNPDTGNSVSEVPRTYQYNKQ 120
QY 121 ERQQYCREFOQKAMTAGOKQBIYGTACRQPDGRWQVISTEK 162
DB 121 ERQQYCREFOQKAMTAGOKQBIYGTACRQPDGRWQVISTEK 162

RESULT 4
AAB81127
ID AAB81127 standard; Protein; 161 AA.
XX
AC AAB81127;
XX
DT 11-JUL-2001 (first entry)
XX
DE Optimised *OspA* protein 17E2 amino acid sequence.

XX Poikilothermic fish; *Piscirickettsia salmonis*; rickettsial pathogen;
KW vaccine; *OspA*; salmonid rickettsial septicemia; rickettsial disease;
KW SRS.
XX
OS *Piscirickettsia salmonis*.
XX Synthetic.
XX

FH Key Location/Qualifiers
FT Region 109..128
FT /label= B_cell_epitope
XX
XX CA2281913-A1.
XX
XX
PD 17-MAR-2001.
XX
XX 17-SEP-1999; 99CA-2281913.
XX
XX 17-SEP-1999; 99CA-2281913.
XX

XX (KAYW/) KAY W W.
PA (BURI/) BURIAN J.
PA (KUYZ/) KUYZK M A.
XX
PI Kay WW, Burian J, Kuzyk MA;
XX
XX WPI; 2001-316844/34.
DR N-PSDB; AAF86247.
XX

PT Method for protecting poikilothermic fish against salmonid rickettsial
PT septicemia and other rickettsial diseases comprises administering a
XX vaccine containing the *OspA* protein of *Piscirickettsia salmonis* -
XX Disclosure; Fig 5; 35pp; English.

CC This invention relates to a method for the protection against infection
CC of a poikilothermic fish by the bacterial pathogen, *Piscirickettsia*
CC salmonis. The method comprises administering an immunogenic amount of a
CC *P. salmonis* specific antigen termed *OspA*, or an immunogenic fragment of
CC *OspA* in the form of a vaccine. The method is used for protecting animals,
CC particularly poikilothermic fish, against the bacterial pathogen
CC *P. salmonis*. The method is also useful for protecting against salmonid
CC rickettsial septicemia (SRS) and other rickettsial diseases. The present
CC sequence represents optimised *P. salmonis* *OspA* protein 17E2. The DNA
CC encoding *OspA* 17E2 (AAF86247) has been optimised for expression in
CC *Escherichia coli*. An *OspA* protein with an N-terminal fusion partner is
CC used in a vaccine to create an anti-*OspA* antibody response.

XX Sequence 161 AA;

Query Match 97.5%; Score 815; DB 22; Length 161;
Best Local Similarity 98.8%; Pred. No. 6.3e-78;
Matches 158; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 3 RGCLOGSSLLIIISVFLVGCQNFSRQEVGAATGAVVGVAGOLFSGSRVMAIGA 62
DB 2 RGCLOGSSLLIIISVFLVGCQNFSRQEVGAATGAVVGVAGOLFSGSRVMAIGA 61
QY 63 GGLIGSKIGSMDOODKIKLNQSLKVKAGQVTRWRNPDTGNSVSEVPRTYQYNKQ 122
DB 62 GGLIGSKIGSMDOODKIKLNQSLKVKAGQVTRWRNPDTGNSVSEVPRTYQYNKQ 121
QY 123 RQQYCREFOQKAMTAGOKQBIYGTACRQPDGRWQVISTEK 162
DB 122 RQQYCREFOQKAMTAGOKQBIYGTACRQPDGRWQVISTEK 161

RESULT 5
AAU97868
ID AAU97868 standard; Protein; 161 AA.
XX
AC AAU97868;
XX
DT 12-AUG-2002 (first entry)
XX
DE *Escherichia coli* codon optimised *OspA*, 17e2.

XX Outer surface lipoprotein; *OspA*; antibacterial; immunosuppressive;
KW vaccine; poikilothermic fish; fin-fish; Rickettsial septicemia;
KW Rickettsial disease; 17e2.

XX *Piscirickettsia salmonis*.
OS Synthetic.
XX
XX CA2339327-A1.
XX
XX 15-MAR-2002.
XX
XX 19-MAR-2001; 2001CA-2339327.
XX
XX 15-SEP-2000; 2000US-0677374.

XX (THOR/) THORNTON J C.
PA (KAYW/) KAY W W.
PA (BURI/) BURIAN J.
PA (KUYZ/) KUYZK M A.
XX
PI Thornton JC, Kay WW, Burian J, Kuzyk MA;
XX
XX WPI; 2002-455221/49.
DR N-PSDB; ABK52402.
XX

PT Inducing immunity in fin fish to Rickettsial septicaemia, comprises
PT administration of an outer surface lipoprotein (OspA) of a bacterial
XX strain, as a vaccine -
XX Example 4; Fig 5; 55pp; English.
XX
XX The invention describes a method of protecting a poikilothermic fish
CC against infection by the bacterial pathogen *Piscirickettsia salmonis*
CC comprising administering either intraperitoneally, by immersion or
CC orally, an immunogenic amount of principal antigen, the OspA (outer
CC surface lipoprotein), its variants, non-lipidated form or antigenic
CC peptides derived or synthesized with or without an adjuvant. The new
CC method is used to provide an outer surface lipoprotein (OspA) of
CC bacterial strain *Piscirickettsia salmonis* as a vaccine to induce immunity
CC in fin-fish against Rickettsial septicaemia and other related
CC Rickettsial diseases caused by either a virus, bacteria or parasite.
CC This is the amino acid sequence of the *Escherichia coli* codon optimised
CC outer surface lipoprotein OspA (17e2) used in the creation of the vaccine
CC described in the invention.
XX Sequence 161 AA;
XX
XX Query Match 97.5%; Score 815; DB 23; Length 161;
XX Best Local Similarity 98.8%; Pred. No. 6.3e-78;
XX Matches 158; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 3 RGLQSSLIISVFLVGCQAQNFPSRQEVGAATGAVVGGVAGQLFGKSGRVAMATGGAVL 62
XX DB 2 RGLQSSLIISVFLVGCQAQNFPSRQEVGAATGAVVGGVAGQLFGKSGRVSMATGGAVL 61
XX
XX QY 63 GGLIGSKIGQSMDDQDKIKLQSLKVKAGQVTRWRNPDTGNSYVPEVRTYQRYNKKQR 122
XX DB 62 GGLIGSKIGQSMDDQDKIKLQSLKVKAGQVTRWRNPDTGNSYVPEVRTYQRYNKKQR 121
XX
XX QY 123 RQCYCFEQQKAMIAQKQEIYGTACRQPDGRWQVISTEK 162
XX DB 122 RQCYCFEQQKAMIAQKQEIYGTACRQPDGRWQVISTEK 161
XX
XX RESULT 6
XX ID AAB81128 standard; Protein; 256 AA.
XX AC AAB81128;
XX
XX DT 11-JUL-2001 (first entry)
XX
XX DE C17E2 OspA construct with N-terminal fusion partner.
XX
XX KW Poikilothermic fish; *Piscirickettsia salmonis*; rickettsial pathogen;
XX vaccine; OspA; salmonid rickettsial septicaemia; rickettsial disease;
XX SRS; 17E2; fusion construct.
XX
XX OS *Piscirickettsia salmonis*.
XX
XX FH Key Location/Qualifiers
XX FT Region 1..95 /label= Undefined_N-terminal_fusion_partner
XX FT Region 96..256 /label= C17E2 OspA
XX FT /note= "Product of OspA gene optimised for expression in
XX FT *Escherichia coli*"
XX
XX CA2281913-A1.
XX
XX PD 17-MAR-2001.
XX
XX PF 17-SEP-1999; 99CA-2281913.
XX
XX PR 17-SEP-1999; 99CA-2281913.
XX
XX RQ (KAYW/) KAY W W.

PA (BURI/) BURIAN J.
PA (KUYZ/) KUYZ M A.
XX
XX Kay WW, Burian J, Kuzyk MA;
XX
XX WPI; 2001-316844/34.
XX N-PSDB; AAF86248.
XX
XX Method for protecting poikilothermic fish against salmonid rickettsial
XX septicaemia and other rickettsial diseases comprises administering a
XX vaccine containing the OspA protein of *Piscirickettsia salmonis* -
XX Example 4; Fig 5; 35pp; English.
XX
XX This invention relates to a method for the protection against infection
XX of a poikilothermic fish by the bacterial pathogen, *Piscirickettsia*
XX *salmonis*. The method comprises administering an immunogenic amount of a
XX *P. salmonis* specific antigen termed OspA, or an immunogenic fragment of
XX OspA in the form of a vaccine. The method is used for protecting animals,
XX particularly poikilothermic fish, against the bacterial pathogen
XX *P. salmonis*. The method is also useful for protecting against salmonid
XX rickettsial septicaemia (SRS) and other rickettsial diseases. The present
XX sequence represents the amino acid sequence of C17E2, a *P. salmonis* OspA
XX construct optimised for expression in *Escherichia coli*, fused to an
XX undefined N-terminal fusion partner. The fusion protein is used in a
XX vaccine to create an anti-OspA antibody response.
XX Sequence 256 AA;
XX
XX Query Match 97.5%; Score 815; DB 22; Length 256;
XX Best Local Similarity 98.8%; Pred. No. 1.1e-77;
XX Matches 158; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 3 RGLQSSLIISVFLVGCQAQNFPSRQEVGAATGAVVGGVAGQLFGKSGRVAMATGGAVL 62
XX DB 97 RGLQSSLIISVFLVGCQAQNFPSRQEVGAATGAVVGGVAGQLFGKSGRVSMATGGAVL 156
XX
XX QY 63 GGLIGSKIGQSMDDQDKIKLQSLKVKAGQVTRWRNPDTGNSYVPEVRTYQRYNKKQR 122
XX DB 157 GGLIGSKIGQSMDDQDKIKLQSLKVKAGQVTRWRNPDTGNSYVPEVRTYQRYNKKQR 216
XX
XX QY 123 RQCYCFEQQKAMIAQKQEIYGTACRQPDGRWQVISTEK 162
XX DB 217 RQCYCFEQQKAMIAQKQEIYGTACRQPDGRWQVISTEK 256
XX
XX RESULT 7
XX ID AAU97869 standard; Protein; 256 AA.
XX AC AAU97869;
XX
XX DT 12-AUG-2002 (first entry)
XX
XX DE *E. coli* codon optimised OspA, 17e2 with N-terminal fusion peptide.
XX
XX KW Outer surface lipoprotein; OspA; antibacterial; immunosuppressive;
XX vaccine; poikilothermic fish; fin-fish; Rickettsial septicaemia;
XX Rickettsial disease; 17e2.
XX
XX OS *Piscirickettsia salmonis*.
XX Synthetic.
XX
XX PN CA2339327-A1.
XX
XX PD 15-MAR-2002.
XX
XX PF 19-MAR-2001; 2001CA-2339327.
XX
XX PR 15-SEP-2000; 2000US-0677374.
XX
XX (THOR/) THORNTON J C.
XX (KAYW/) KAY W W.

PA (BURI/) BURIAN J.
XX (KUYZ/) KUYZK M A.
PI Thornton JC, Kay WW, Burian J, Kuzyk MA;
XX WPI; 2002-455221/49.
DR N-PSDB; ABK52403.
XX
PT Inducing immunity in fin fish to Rickettsial septicaemia, comprises
PT administration of an outer surface lipoprotein (OspA) of a bacterial
PT strain, as a vaccine -
XX
PS Example 4; Fig 5; 55pp; English.
XX
CC The invention describes a method of protecting a poikilothermic fish
CC against infection by the bacterial pathogen *Piscirickettsia salmonis*
CC comprising administering either intraperitoneally, by immersion or
CC orally, an immunogenic amount of principal antigen, the OspA (outer
CC surface lipoprotein), its variants, non-lipidated form or antigenic
CC peptides derived or synthesized with or without an adjuvant. The new
CC method is used to provide an outer surface lipoprotein (OspA) of
CC bacterial strain *Piscirickettsia salmonis* as a vaccine to induce immunity
CC in fin-fish against Rickettsial septicaemia and other related
CC Rickettsial diseases caused by either a virus, bacteria or parasite.
CC This is the amino acid sequence of the *Escherichia coli* codon optimised
CC outer surface lipoprotein OspA (17e2) with an N-terminal fusion used in
CC the creation of the vaccine described in the invention.
XX
SQ Sequence 256 AA;

Query Match 97.5%; Score 815; DB 23; Length 256;
Best Local Similarity 98.8%; Pred. No. 1.1e-77;
Matches 158; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 RCLQGSSLIISVFLVGCQAQNFQSGVGAATGAVVGGVAGQLFGKSGRVAIGAVAL 62
Db 97 RCLQGSSLIISVFLVGCQAQNFQSGVGAATGAVVGGVAGQLFGKSGRVAIGAVAL 156

QY 63 GGLIGSKIGSQMDQDKIKLNQSLKVKAGQVTRWRNPDTGNSYVPEVRTYQRYNKOER 122
Db 157 GGLIGSKIGSQMDQDKIKLNQSLKVKAGQVTRWRNPDTGNSYVPEVRTYQRYNKOER 216

QY 123 RQCYCREFOQKAMIAQKQEIYGTACRQPDGRWQVISTEK 162
Db 217 RQCYCREFOQKAMIAQKQEIYGTACRQPDGRWQVISTEK 256

RESULT 8
ABJ18820
ID ABJ18820 standard; Protein; 182 AA.
XX
AC ABJ18820;
XX
DT 27-FEB-2003 (first entry)
XX
DE Pseudomonas aeruginosa biofilm formation-related protein #84.
XX
KW Biofilm formation modulation; biofilm-associated disease;
KW cystic fibrosis; AIDS; middle ear infection; acne; periodontal disease;
KW catheter-associated infection; medical device-associated infection.
XX
OS Pseudomonas aeruginosa.
XX
PN WO200285295-A2.
XX
PD 31-OCT-2002.
XX
PF 19-APR-2002; 2002WO-US12532.
XX
PR 20-APR-2001; 2001US-285190P.
PR 24-OCT-2001; 2001US-344142P.
XX
PA (IOWA) UNIV IOWA RES FOUND.

PA (HARD) HARVARD COLLEGE.
XX
PI Whiteley M, Bangera MG, Lory S, Greenberg EP;
XX WPI; 2003-075601/07.
DR N-PSDB; ABT14642.
XX
PT Identifying compound capable of modulating biofilm formation by
PT bacteria/bacterial antibiotic resistance, useful for treatment of
PT biofilm associated disease -
XX
PS Claim 4; Page 152; 154pp; English.
XX
CC The invention comprises a method for identifying a compound capable of
CC modulating biofilm formation by bacteria. The method of the invention is
CC useful for identifying a compound capable of modulating biofilm formation
CC by bacteria or modulating bacterial antibiotic resistance. The method of
CC the invention is also useful for diagnosing and treating a subject
CC (especially an immunocompromised human) that is afflicted with a biofilm-
CC associated disease or disorder, such as: cystic fibrosis; AIDS; middle
CC ear infections; acne; periodontal disease; catheter-associated
CC infections; and medical device-associated infections. The present amino
CC acid sequence represents a protein that is used in the invention.
XX
SQ Sequence 182 AA;

Query Match 15.4%; Score 128.5; DB 24; Length 182;
Best Local Similarity 34.8%; Pred. No. 2.1e-05;
Matches 32; Conservative 13; Mismatches 28; Indels 19; Gaps 2;

QY 28 QEVGAATGAVVGGVAGQLFGKSGRVAIGAVALGIGSKIGSQMDQDKI----- 80
Db 70 QIAGTAIGAVVGGLLGNIGGGTGKTIATVAGVGGVAGNKVQSGMQERDPTTTTETRC 129

QY 81 -KLNQSLKVK-----KAGQVTRWRNP 100
Db 130 STVHDSSEKVVGYDYVKYMLDGGKAGQIRMERDP 161

RESULT 9
AAB81130
ID AAB81130 standard; Peptide; 20 AA.
XX
AC AAB81130;
XX
DT 11-JUL-2001 (first entry)
XX
DE OspA B-cell epitope peptide #2.
XX
KW Poikilothermic fish; *Piscirickettsia salmonis*; rickettsial pathogen;
KW vaccine; OspA; salmonid rickettsial septicaemia; rickettsial disease;
KW SRS; antibody.
XX
OS *Piscirickettsia salmonis*.
XX
PN CA2281913-A1.
XX
PD 17-WAR-2001.
XX
PF 17-SEP-1999; 99CA-2281913.
XX
PR 17-SEP-1999; 99CA-2281913.
XX
PA (KAYW/) KAY W W.
PA (BURI/) BURIAN J.
PA (KUYZ/) KUYZK M A.
XX
PI Kay WW, Burian J, Kuzyk MA;
XX
DR WPI; 2001-316844/34.
XX
PT Method for protecting poikilothermic fish against salmonid rickettsial
PT septicaemia and other rickettsial diseases comprises administering a

PT vaccine containing the Ospa protein of *Piscirickettsia salmonis* -
 XX Example 2; Page 17; 35pp; English.
 XX This invention relates to a method for the protection against infection
 CC of a poikilothermic fish by the bacterial pathogen, *Piscirickettsia*
 CC salmonis. The method comprises administering an immunogenic amount of a
 CC *P. salmonis* specific antigen termed Ospa, or an immunogenic fragment of
 CC Ospa in the form of a vaccine. The method is used for protecting animals,
 CC particularly poikilothermic fish, against the bacterial pathogen
 CC *P. salmonis*. The method is also useful for protecting against salmonid
 CC rickettsial septicaemia (SR) and other rickettsial diseases. The present
 CC sequence represents an immunogenic epitope of the *P. salmonis* Ospa
 CC protein. The peptide is used to raise rabbit anti-Ospa antibodies.
 XX
 SQ Sequence 20 AA;
 Query Match 13.4%; Score 112; DB 22; Length 20;
 Best Local Similarity 100.0%; Pred. No. 6.9e-05;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 110 PVRTYQRYNKQERRQCYCRE 129
 DB 1 PVRTYQRYNKQERRQCYCRE 20
 RESULT 10
 AAU97871
 ID AAU97871 standard; Peptide; 20 AA.
 XX
 AC AAU97871;
 XX
 DT 12-AUG-2002 (first entry)
 XX
 DE Outer surface lipoprotein Ospa based peptide #2.
 XX
 KW Outer surface lipoprotein; Ospa; antibacterial; immunosuppressive;
 KW vaccine; poikilothermic fish; fin-fish; Rickettsial septicaemia;
 KW Rickettsial disease.
 XX
 OS *Piscirickettsia salmonis*.
 OS Synthetic.
 XX
 PN CA2339327-A1.
 XX
 PD 15-MAR-2002.
 XX
 PF 19-MAR-2001; 2001CA-2339327.
 XX
 PR 15-SEP-2000; 2000US-0677374.
 XX
 PA (THOR/) THORNTON J C.
 PA (KAYW/) KAY W W.
 PA (BURI/) BURIAN J.
 PA (KUZY/) KUZYSK M A.
 XX
 PI Thornton JC, Kay WW, Burian J, Kuzysk MA;
 XX WPI; 2002-455221/49.
 DR
 XX Inducing immunity in fin fish to Rickettsial septicaemia, comprises
 PT administration of an outer surface lipoprotein (Ospa) of a bacterial
 PT strain, as a vaccine -
 PT
 XX Example 2; Page 17; 55pp; English.
 PS
 XX The invention describes a method of protecting a poikilothermic fish
 CC against infection by the bacterial pathogen *Piscirickettsia salmonis*
 CC comprising administering either intraperitoneally, by immersion or
 CC orally, an immunogenic amount of principal antigen the Ospa (outer
 CC surface lipoprotein), its variants, non-lipidated form or antigenic
 CC peptides derived or synthesized with or without an adjuvant. The new
 CC method is used to provide an outer surface lipoprotein (Ospa) of

CC bacterial strain *Piscirickettsia salmonis* as a vaccine to induce immunity
 CC in fin-fish against Rickettsial septicaemia and other related
 CC Rickettsial diseases caused by either a virus, bacteria or parasite.
 CC This sequence represents a synthetic peptide used to create polyclonal
 CC rabbit antibodies against the *Piscirickettsia salmonis* outer surface
 CC lipoprotein, Ospa.
 XX
 SQ Sequence 20 AA;
 Query Match 13.4%; Score 112; DB 23; Length 20;
 Best Local Similarity 100.0%; Pred. No. 6.9e-05;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 110 PVRTYQRYNKQERRQCYCRE 129
 DB 1 PVRTYQRYNKQERRQCYCRE 20
 RESULT 11
 AA017565
 ID AA017565 standard; Protein; 197 AA.
 XX
 AC AA017565;
 XX
 DT 19-JUL-2002 (first entry)
 XX
 DE M catarrhalis MCA100414 protein SEQ ID NO: 10.
 XX
 KW Moraxella; vaccine; respiratory tract infection; antiinflammatory;
 KW auditory; antibacterial; otitis media; sinusitis; pneumonia.
 XX
 OS Moraxella catarrhalis.
 XX
 PN WO200218595-A2.
 XX
 PD 07-MAR-2002.
 XX
 PF 28-AUG-2001; 2001WO-CA01221.
 XX
 PR 28-AUG-2000; 2000US-228294P.
 PR 28-AUG-2000; 2000US-228295P.
 PR 28-AUG-2000; 2000US-228296P.
 PR 29-AUG-2000; 2000US-228438P.
 PR 29-AUG-2000; 2000US-228439P.
 PR 29-AUG-2000; 2000US-228440P.
 PR 29-AUG-2000; 2000US-228441P.
 PR 29-AUG-2000; 2000US-228442P.
 PR 29-AUG-2000; 2000US-228443P.
 PR 29-AUG-2000; 2000US-228511P.
 PR 29-AUG-2000; 2000US-228512P.
 PR 29-AUG-2000; 2000US-228742P.
 PR 29-AUG-2000; 2000US-228773P.
 PR 01-SEP-2000; 2000US-229465P.
 PR 01-SEP-2000; 2000US-229474P.
 PR 01-SEP-2000; 2000US-229475P.
 PR 01-SEP-2000; 2000US-229478P.
 PR 05-SEP-2000; 2000US-229740P.
 PR 05-SEP-2000; 2000US-229803P.
 PR 05-SEP-2000; 2000US-229804P.
 PR 05-SEP-2000; 2000US-229805P.
 PR 05-SEP-2000; 2000US-229806P.
 PR 05-SEP-2000; 2000US-229809P.
 PR 05-SEP-2000; 2000US-229811P.
 PR 06-SEP-2000; 2000US-230214P.
 PR 06-SEP-2000; 2000US-230250P.
 PR 06-SEP-2000; 2000US-230252P.
 XX
 XX (AVET) AVENTIS PASTEUR LTD.
 XX
 XX Loosmore S, Wang J, Bradley B, Ochs M, Yang Y;
 XX WPI; 2002-401721/43.
 DR N-PSDB; AAL46497.

XX Moraxella polypeptide and polynucleotides useful as vaccine for
PT immunizing a host e.g. humans against disease e.g. otitis media,
PT pneumonia, caused by infection of the bacteria -
XX PS Claim 28; Fig 9; 277pp; English.
XX The present invention provides the protein and coding sequences of
CC proteins from Moraxella catarrhalis. These can be used to produce
CC vaccines which protect against M. catarrhalis infection, which can cause
CC otitis media, respiratory infection, sinusitis, and pneumonia. The
CC present sequence is a protein of the invention.
XX SQ Sequence 197 AA;
Query Match 12.9%; Score 108; DB 23; Length 197;
Best Local Similarity 32.4%; Pred. No. 0.0033;
Matches 36; Conservative 16; Mismatches 35; Indels 24; Gaps 5;
QY 8 GSSLIISVFLVGCANFROEVGAATG-----AVVGVAGOLFPGKSGRVAMAIGG 59
DB 7 GVLLASSMALAGCANTGT---TGNGTGFGANVKNVAVAGAL---GGTAISKATGG 60
QY 60 -----AVLGGILGSKIGSQMDQDKIKLQSLKVKAGQVTRWRNPDTGN 104
DB 61 EKTGRDAILGAAGAAGAYMERQAK----QIEQQMQGTGVTVTHDITDGN 107
RESULT 12
AAB20105
ID AAB20105 standard; Protein; 224 AA.
XX AC AAB20105;
XX DT 23-APR-2001 (first entry)
XX DE Moraxella catarrhalis BASB113 protein.
XX KW BASB113; infection; otitis media; pneumonia; therapy; diagnosis;
XX KW antibacterial; antimicrobial.
XX OS Moraxella catarrhalis.
XX PN WO200100836-A1.
XX PD 04-JAN-2001.
XX PF 23-JUN-2000; 2000WO-EP05951.
XX PR 25-JUN-1999; 99GB-0015044.
XX PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX PI Thonnard J;
XX DR WPI; 2001-112458/12.
XX DR N-PSDB; AAF30043.
XX PT New BASB113 polypeptide isolated from Moraxella catarrhalis bacterium,
PT useful for diagnosing and producing vaccines against bacterial
PT infections such as otitis media and pneumonia -
XX PS Claim 1; Page 67; 86pp; English.
XX The present sequence is that of BASB113 protein from Moraxella
CC catarrhalis strain Mc2931 (ATCC 43617), a causative agent of otitis
CC media in children and pneumonia in adults. The invention provides
CC BASB113 polypeptides, and polynucleotides encoding them, as well as
CC expression vectors, host cells and methods for producing BASB113
CC polypeptides using recombinant methods. Also claimed is a vaccine
CC composition comprising a BASB113 polypeptide, an immunogenic
CC fragment of a BASB113 polypeptide, or a polypeptide having at least
CC 85% amino acid sequence identity to BASB113, or comprising a

CC polynucleotide encoding such a polypeptide. A claimed method of
CC diagnosing a Moraxella infection involves identifying a BASB113
CC polypeptide or antibody. A claimed therapeutic composition useful
CC in treating humans with M. catarrhalis infection comprises at least
CC 1 antibody directed against a BASB113 polypeptide. BASB113
CC polypeptides also have utility in raising specific antibodies,
CC and in screening for antibacterial drugs.
XX SQ Sequence 224 AA;
Query Match 12.9%; Score 108; DB 22; Length 224;
Best Local Similarity 32.4%; Pred. No. 0.0039;
Matches 36; Conservative 16; Mismatches 35; Indels 24; Gaps 5;
QY 8 GSSLIISVFLVGCANFROEVGAATG-----AVVGVAGOLFPGKSGRVAMAIGG 59
DB 7 GVLLASSMALAGCANTGT---TGNGTGFGANVKNVAVAGAL---GGTAISKATGG 60
QY 60 -----AVLGGILGSKIGSQMDQDKIKLQSLKVKAGQVTRWRNPDTGN 104
DB 61 EKTGRDAILGAAGAAGAYMERQAK----QIEQQMQGTGVTVTHDITDGN 107
RESULT 13
ABG80432
ID ABG80432 standard; Protein; 224 AA.
XX AC ABG80432;
XX DT 29-NOV-2002 (first entry)
XX DE Moraxella catarrhalis PAL-2 protein.
XX KW Hyperblebbing; Gram-negative bacterium; genetically modified; Tol gene;
KW peptidoglycan-binding; peptidoglycan-associated slice; outer-membrane;
KW bacterial infection; vesicle-shedding; Bleb; filter sterilised;
KW detergent; deoxycholate; homogeneity; antibacterial; vaccine;
XX KW TolQ; Tola; OmpCD; xOmpA; PAL-1; PAL-2.
XX OS Moraxella catarrhalis.
XX PN WO200262378-A2.
XX PD 15-AUG-2002.
XX PF 08-FEB-2002; 2002WO-EP01361.
XX PR 08-FEB-2001; 2001GB-0003171.
XX PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX PI Berthet FJ, Denoel P, Neyt CA, Poolman J, Thonnard J;
XX DR WPI; 2002-657509/70.
XX DR N-PSDB; ABS66202.
XX PT Novel hyperblebbing Gram-negative bacterium that has been genetically
PT modified by down-regulating expression of Tol genes, and/or attenuating
PT peptidoglycan-binding activity useful for treating bacterial infection
XX PS Disclosure; Page 71; 71pp; English.
XX The present invention relates to a new hyperblebbing Gram-negative
CC bacterium genetically modified by one or more processes selected from
CC down-regulating expression of one or more Tol genes and attenuating the
CC peptidoglycan-binding activity by mutation of one or more gene(s)
CC encoding a protein comprising a peptidoglycan-associated site. The
CC invention is useful in a method of treatment of the human or animal
CC body. The invention is also useful for protecting an individual against
CC a bacterial infection. The invention has improved outer-membrane
CC vesicle-shedding properties. Blebs are more easily made in higher yield
CC from the invention, and are more homogeneous in nature and can be more

CC readily filter sterilised. The blebs can be made and harvested without
 CC the use of detergents such as deoxycholate, thus obviating chromatography
 CC purification and ultra centrifugation steps. Vesicles prepared from the
 CC invention have reduced particle size (allowing sterile filtration through
 CC 0.22 µm pores), increased batch homogeneity, and a superior yield.
 CC The present amino acid sequence represents a Moraxella catarrhalis
 CC protein, as described in the invention.

XX
 XX SQ Sequence 224 AA;
 Query Match 12.9%; Score 108; DB 23; Length 224;
 Best Local Similarity 32.4%; Pred. No. 0.0039;
 Matches 36; Conservative 16; Mismatches 35; Indels 24; Gaps 5;
 QY 8 GSSLIIISVFLVGCQNFRQEVGAATG-----AVVGGVAGOLFQKSGRVRMAIGG 59
 Db 7 GVVLASSWALGACANTGT---TGNGTGGGANVKNVAVAGAL---GGTAISKATGG 60
 QY 60 -----AVLGGIIGSKIGQSMQDDKIKLNQSLKVKAGQVTRWRNPDTGN 104
 Db 61 EXTGRDAILGAAGVGAAGAYMERQAK---QIEQQWQGTGVTVTHDTDGN 107

RESULT 14
 AAY34487
 ID AAY34487 standard; Protein; 223 AA.

XX AC AAY34487;
 XX
 DT 20-MAR-2003 (updated)
 DT 25-AUG-1999 (first entry)
 XX
 DE Porphorymonas gingivalis protein PG3.
 XX
 KW Porphorymonas gingivalis; PG; periodontal disease; gingivitis;
 KW vaccine; antigenic.
 XX
 OS Porphorymonas gingivalis.
 XX
 PN WO9929870-A1.
 XX
 PD 17-JUN-1999.

XX PF 10-DEC-1998; 98WO-AU01023.
 XX
 PR 10-DEC-1997; 97AU-0000839.
 PR 31-DEC-1997; 97AU-0001182.
 PR 30-JAN-1998; 98AU-0001546.
 PR 10-MAR-1998; 98AU-0002264.
 PR 09-APR-1998; 98AU-0002911.
 PR 23-APR-1998; 98AU-0003128.
 PR 05-MAY-1998; 98AU-0003338.
 PR 22-MAY-1998; 98AU-0003654.
 PR 29-JUL-1998; 98AU-0004917.
 PR 30-JUL-1998; 98AU-0004963.
 PR 04-AUG-1998; 98AU-0005028.
 XX
 PA (CSLC-) CSL LTD.

XX PI Agius CT, Barr IG, Hocking DM, Margetts MB, Patterson MA;
 PI Ross BC, Rothe LJ, Webb EA;
 XX
 DR WPI; 1999-385613/32.
 DR N-PSDB; AAX91705.

XX PT Antigenic Porphorymonas gingivalis peptides for preventing
 PT gingivitis
 XX
 PS Claim 1; Page 469; 588pp; English.

XX
 CC AAX91536 to AAX91801 encode two hundred and sixty six antigenic
 CC Porphorymonas gingivalis (PG) polypeptide sequences given in AAX34318 to
 CC AAX34583. AAX91802 to AAX91989 represent PCR primers used in the

CC isolation of the PG polypeptides. The PG polypeptides have antibacterial
 CC activity with a vaccine mechanism of action. The PG polypeptides can be
 CC used as vaccines especially against Porphorymonas gingivalis. Probes can
 CC be used to detect Porphorymonas gingivalis in standard hybridisation
 CC assays. Porphorymonas gingivalis is involved in periodontal disease
 CC especially gingivitis.
 CC (Updated on 20-MAR-2003 to correct PR field.)

XX
 XX SQ Sequence 223 AA;
 Query Match 12.6%; Score 105.5; DB 20; Length 223;
 Best Local Similarity 36.1%; Pred. No. 0.0072;
 Matches 26; Conservative 12; Mismatches 31; Indels 3; Gaps 2;
 QY 9 SSSLIIISVFLVGCQNFRQEVGAATGAVVGGVAGOLFQKSGRVRMAIGAVLGGIIG 67
 Db 8 ASVLAVALVFAGCGLN--NMAKGGIIGAGVGAIGAGVGNVAGNTAVGAIVGTAVGAAG 65
 QY 68 SKIGQSMQDDK 79
 Db 66 ALIGKKMDKQKK 77

RESULT 15
 AAY34362
 ID AAY34362 standard; Protein; 230 AA.

XX AC AAY34362;
 XX
 DT 20-MAR-2003 (updated)
 DT 25-AUG-1999 (first entry)
 XX
 DE Porphorymonas gingivalis protein PG3.
 XX
 KW Porphorymonas gingivalis; PG; periodontal disease; gingivitis;
 KW vaccine; antigenic.
 XX
 OS Porphorymonas gingivalis.
 XX
 PN WO9929870-A1.
 XX
 PD 17-JUN-1999.

XX PF 10-DEC-1998; 98WO-AU01023.
 XX
 PR 10-DEC-1997; 97AU-0000839.
 PR 31-DEC-1997; 97AU-0001182.
 PR 30-JAN-1998; 98AU-0001546.
 PR 10-MAR-1998; 98AU-0002264.
 PR 09-APR-1998; 98AU-0002911.
 PR 23-APR-1998; 98AU-0003128.
 PR 05-MAY-1998; 98AU-0003338.
 PR 22-MAY-1998; 98AU-0003654.
 PR 29-JUL-1998; 98AU-0004917.
 PR 30-JUL-1998; 98AU-0004963.
 PR 04-AUG-1998; 98AU-0005028.
 XX
 PA (CSLC-) CSL LTD.

XX PI Agius CT, Barr IG, Hocking DM, Margetts MB, Patterson MA;
 PI Ross BC, Rothe LJ, Webb EA;
 XX
 DR WPI; 1999-385613/32.
 DR N-PSDB; AAX91580.

XX PT Antigenic Porphorymonas gingivalis peptides for preventing
 PT gingivitis
 XX
 PS Claim 1; Page 325-326; 588pp; English.

XX
 CC AAX91536 to AAX91801 encode two hundred and sixty six antigenic
 CC Porphorymonas gingivalis (PG) polypeptide sequences given in AAX34318 to
 CC AAX34583. AAX91802 to AAX91989 represent PCR primers used in the

CC isolation of the PG polypeptides. The PG polypeptides have antibacterial
CC activity with a vaccine mechanism of action. The PG polypeptides can be
CC used as vaccines especially against Porphyromonas gingivalis. Probes can
CC be used to detect Porphyromonas gingivalis in standard hybridisation
CC assays. Porphyromonas gingivalis is involved in periodontal disease
CC especially gingivitis.
CC (Updated on 20-MAR-2003 to correct PR field.)
XX

SQ Sequence 230 AA;
Query Match 12.6%; Score 105.5; DB 20; Length 230;
Best Local Similarity 36.1%; Pred.No. 0.0074;
Matches 26; Conservative 12; Mismatches 31; Indels 3; Gaps 2;
QY 9 SSLIIISVFLVGCANFSRQEVGAATGAVVGGVAGQLFGKGSGRVAM-AIGGAVLGGLIG 67
Db 15 ASVLAVALVFAGCGLN--NMAKGGLIGAGVGAIGAGVGNVAGNTAVGAIVGTAVGGAG 72
QY 68 SKIGSQMDQDDK 79
Db 73 ALIGKKMDKQKK 84

Search completed: November 5, 2003, 20:12:49
Job time : 30.3782 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 20:11:02 ; Search time 10.3523 Seconds
(without alignments)
662.108 Million cell updates/sec

Title: US-09-677-374-2

Perfect score: 836

Sequence: 1 MNRGCLQSSLIILISFLVG.....IYGTAQPDGRWQVISTEK 162

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/1/1aa/5A-COMB.pep.*
- 2: /cgn2_6/ptodata/1/1aa/5B-COMB.pep.*
- 3: /cgn2_6/ptodata/1/1aa/6A-COMB.pep.*
- 4: /cgn2_6/ptodata/1/1aa/6B-COMB.pep.*
- 5: /cgn2_6/ptodata/1/1aa/PCTUS-COMB.pep.*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	128.5	15.4	192	4	US-09-252-991A-22817
2	116	13.9	165	4	US-09-252-991A-19701
3	101.5	12.1	306	4	US-09-252-991A-21754
4	98.5	11.8	387	4	US-09-252-991A-23334
5	93.5	11.2	112	4	US-09-252-991A-31960
6	90	10.8	126	4	US-09-328-352-8009
7	89	10.6	199	4	US-09-328-352-7802
8	87.5	10.5	215	4	US-09-328-352-6750
9	85.5	10.2	730	4	US-09-328-352-4442
10	85	10.2	258	4	US-09-328-352-4253
11	83.5	10.0	127	4	US-09-252-991A-28397
12	83	9.9	1034	4	US-09-252-991A-26658
13	82	9.8	734	4	US-09-252-991A-33036
14	81.5	9.7	217	4	US-09-328-352-7068
15	81.5	9.7	651	3	US-08-556-978B-19
16	81.5	9.7	651	1	US-09-247-806-1
17	81.5	9.7	718	1	US-08-425-069-2
18	81.5	9.7	718	2	US-08-317-844B-2
19	81.5	9.7	747	3	US-09-034-177-3
20	81	9.7	387	4	US-09-328-352-6442
21	80.5	9.6	84	4	US-09-107-532A-5149
22	79.5	9.5	150	4	US-09-328-352-6423
23	79	9.4	572	4	US-09-252-991A-23878
24	78.5	9.4	104	4	US-09-107-532A-7222
25	78.5	9.4	172	4	US-09-252-991A-27926
26	78	9.3	865	1	US-07-803-633A-13
27	77.5	9.3	255	4	US-09-553-498-8

28	77.5	9.3	255	4	US-09-618-869-8	Sequence 8, Appli
29	77	9.2	141	4	US-09-328-352-7100	Sequence 7100, Ap
30	77	9.2	551	2	US-09-067-351-2	Sequence 2, Appli
31	77	9.2	551	3	US-09-360-490-2	Sequence 2, Appli
32	76.5	9.2	259	4	US-09-328-352-5775	Sequence 5775, Ap
33	76.5	9.2	518	4	US-09-252-991A-19162	Sequence 19162, A
34	76.5	9.2	849	4	US-09-252-991A-31525	Sequence 31525, A
35	76.5	9.2	1415	4	US-09-252-991A-26438	Sequence 26438, A
36	76.5	9.2	2516	3	US-08-374-077C-2	Sequence 2, Appli
37	76.5	9.2	2516	3	US-08-895-590-2	Sequence 2, Appli
38	76.5	9.2	2516	4	US-09-539-879A-2	Sequence 2, Appli
39	76	9.1	397	4	US-09-252-991A-30059	Sequence 30059, A
40	76	9.1	1021	1	US-07-910-760-12	Sequence 12, Appl
41	76	9.1	1021	1	US-08-440-519-12	Sequence 12, Appl
42	76	9.1	1021	4	US-08-440-549-12	Sequence 12, Appl
43	75.5	9.0	82	4	US-09-107-532A-4748	Sequence 4748, Ap
44	75.5	9.0	204	4	US-09-107-532A-6584	Sequence 6584, Ap
45	75	9.0	141	4	US-09-252-991A-23427	Sequence 23427, A

ALIGNMENTS

RESULT 1
US-09-252-991A-22817
; Sequence 22817, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: 107196.136
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22817
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22817

Query Match 15.4%; Score 128.5; DB 4; Length 192;
Best Local Similarity 34.8%; Pred. No. 1.3e-06;
Matches 32; Conservative 13; Mismatches 28; Indels 19; Gaps 2;

QY 28 QEVGAATGAVVGGVAGQLFGKSGRVAIMAIGAVLGLGSKIGQSMDDQDKI-----80
DB 80 QIAGTAIGAVVGGLLGNQIGGGTGGKATVAGAVGGYAGNVQSGMQERDYYTTTETRC 139

QY 81 -KLNOSLEKV-----KAGQVTRWNP 100
DB 140 STVHDSSEKVGVDVKYMLDGRKQIRMERDP 171

RESULT 2
US-09-252-991A-19701
; Sequence 19701, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: 107196.136
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27

US-09-328-352-6750
Query Match 10.5%; Score 87.5; DB 4; Length 215;
Best Local Similarity 23.5%; Pred. No. 0.057; Mismatches 21; Indels 41; Gaps 7;
Matches 35; Conservative 21; Mismatches 21; Indels 41; Gaps 7;
QY 1 MNRGCLQGSSLIISVFLVGCANF--SROEVGAATGAVVGGVAGQLFGKSGRVAAMAIGA 60
DB 64 MNEG---GENLVATGAGTLG-----GAAVGAAGVVGGP-----PGAUVGGIIGG 105
QY 61 VLGLIGSKIGSQMDQDKIKLNQSLKVKAGQVTRWNPDT--GNSYSVEVPRTYQRYN 118
DB 106 VVGAIGANDIACTNNKQD-----DSNDWQEDNEEDNRYWRENYKKMP-----YY 145
QY 119 KQERQQYCRBFQOQKAMIAQKQEIYGTGTA 147
DB 146 SEDKNLEYDRDY-RAAYRLGYENRVHNA 173

RESULT 9

US-09-328-352-4442
; Sequence 4442, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4442
; LENGTH: 730
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4442

Query Match 10.2%; Score 85.5; DB 4; Length 730;
Best Local Similarity 22.0%; Pred. No. 0.48; Mismatches 33; Conservative 30; Mismatches 30; Indels 35; Gaps 6;
QY 8 GSSLIISVFLVGCQAQ-----NFSROEVGAATG-----AVVGGVAGQ 44
DB 329 GVKLALVLLIGFFRISDIITAGVISNVFYQDLNFSKEQIAEAVKVYGVIFSLVGGFLGG 388
QY 45 LFGKSGRVAAMAIGAVLG-----GLIGSKIGSQMDQDKIKLNQSLKVKAGQVTR 96
DB 389 LLAQRMNIMKLMFVGAIVLASSTNLIFIGLV--KSGKPLDMVD-VQIGNHHYKVPDEVGY 445
QY 97 WRNPDTGNSY-SVEPRTYQRYNKQERROQ 125
DB 446 WKLVFSSFSNSTQIDVKAGYALQDOQAE 475

RESULT 10

US-09-328-352-4253
; Sequence 4253, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4253
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4253

Query Match 10.2%; Score 85; DB 4; Length 258;

US-09-328-352-7802
Query Match 10.8%; Score 90; DB 4; Length 126;
Best Local Similarity 31.8%; Pred. No. 0.015; Mismatches 25; Indels 20; Gaps 4;
Matches 27; Conservative 13; Mismatches 13; Indels 20; Gaps 4;
QY 9 SSSLIIISVFLVGCANF--SROEVGAATGAVVGGVAGQLFGKSGR--VAMAIGGAV---- 61
DB 18 STVILSTVMFTGC-QNMSFSDQRIAGAAAGALGGGLGNHVGIGAGLGAAGVAGVGSNT 76
QY 62 -----LGLIGSKIGSQSM 74
DB 77 QGSGKQTTTSAIGAIGAGIGSWGKAI 101

RESULT 7

US-09-328-352-7802
; Sequence 7802, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7802
; LENGTH: 199
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7802

Query Match 10.6%; Score 89; DB 4; Length 199;
Best Local Similarity 25.7%; Pred. No. 0.035; Mismatches 36; Conservative 19; Mismatches 43; Indels 42; Gaps 6;
QY 35 GAVVGGVAGQLFGKSGRVAAMAIGAVLGLIGSKIGSQMDQ----- 76
DB 32 GGMGLSVLQMGNTS---SGAQGG--LGGVILGSLVGGVGTGNNTPQTGGVQSLLIAVV 86
QY 77 -----QDKIKLNQSLKVK---AGQVTRWRNPDTGNSYSVEPRTYQRYNKQERROQ 125
DB 87 PLLILGWVQQGGQLAALKLGAGLGSQVQSWVDPNQSN--SEVPYQQLQSLFNPADIEQ 144
QY 126 YCREFOQKAMIAQKQEIYGTGTA 145
DB 145 VAQQAQ-----APKEQVYTG 158

RESULT 8

US-09-328-352-6750
; Sequence 6750, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6750
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii

Best Local Similarity 29.7%; Pred. No. 0.14; Mismatches 14; Indels 24; Gaps 2;

Matches 22; Conservative 14; Mismatches 14; Indels 24; Gaps 2;

QY 30 VGAATGAVVGGVAGLFGK-----GSGRVAMAIGGAVLGGIGKIGSQSMDQ 76

Db 166 VSFATAGTGGTGAIGGNFSSGNMSYQATGAGAGGAGGLIVAAIINAEVKGIIIG 225

QY 77 QDKIKLNQSLKVK 90

Db 226 L-PIKESPFMEKLR 238

RESULT 11

US-09-252-991A-28397

; Sequence 28397, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 28397

; LENGTH: 127

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-28397

Query Match 10.0%; Score 83.5; DB 4; Length 127;

Best Local Similarity 32.9%; Pred. No. 0.079; Mismatches 21; Indels 17; Gaps 3;

Matches 23; Conservative 9; Mismatches 21; Indels 17; Gaps 3;

QY 5 CLOGSSLIISVPLVGCANFSRQEVGAATGAVVGGVAGLFGKSGSRVAMAIG-----58

Db 39 CVSGGTL---SGMIVGAVD-----GAATGMAIGKGGAGGFGGALSQVLVLVPTA 88

QY 59 -GAVLGGLIG 67

Db 89 MGAIAGGTG 98

RESULT 12

US-09-252-991A-26658

; Sequence 26658, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 26658

; LENGTH: 1034

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-26658

Query Match 9.9%; Score 83; DB 4; Length 1034;

Best Local Similarity 33.3%; Pred. No. 1.5; Mismatches 29; Indels 29; Gaps 2;

Matches 23; Conservative 11; Mismatches 29; Indels 29; Gaps 2;

QY 22 AQNFSRQEVGAATGAVVGGVAGLFGK-----GSGRVAMAIGGAVLGGIGKIGSQSMDQ 76

Db 866 AHAIISRGDVSQGFNSLVGGLVGHNGHGVNVDASGRVSAASV-GGLVGSNAGSILSA 924

QY 77 QDKIKLNQSLKVK 85

Db 925 RSSSTVNGS 933

RESULT 13

US-09-252-991A-33036

; Sequence 33036, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 33036

; LENGTH: 734

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-33036

Query Match 9.8%; Score 82; DB 4; Length 734;

Best Local Similarity 27.5%; Pred. No. 1.2; Mismatches 28; Indels 20; Gaps 3;

Matches 25; Conservative 18; Mismatches 28; Indels 20; Gaps 3;

QY 35 GAVVGGVAGLFGKSGSRVAMAIGGAVLGGIGKIGSQSMDQDKIKLNQSLKVKAGQV 94

Db 14 GLILGGPAGE-----SLVGALLGGLSGLAQVSLQNLAAQQNRLKQMSBFA 62

QY 95 TRWRNPDTGNSYSVEPVRTYQRYNKQERRQ 125

Db 63 ERFER-----GTEVI--HQRLLRVERQAO 84

RESULT 14

US-09-328-352-7068

; Sequence 7068, Application US/09328352

; Patent No. 6562958

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

; FILE REFERENCE: GTC99-03PA

; CURRENT APPLICATION NUMBER: US/09/328,352

; PRIOR FILING DATE: 1999-06-04

; NUMBER OF SEQ ID NOS: 8252

; SEQ ID NO 7068

; LENGTH: 217

; TYPE: PRT

; ORGANISM: Acinetobacter baumannii

US-09-328-352-7068

Query Match 9.7%; Score 81.5; DB 4; Length 217;

Best Local Similarity 25.2%; Pred. No. 0.27; Mismatches 52; Indels 35; Gaps 4;

Matches 36; Conservative 20; Mismatches 52; Indels 35; Gaps 4;

QY 6 LQSSLIISVPLVGCANFSRQEVGAATGAVVGGVAGLFGKSG-----51

Db 57 LASALIATTSMTVTVAHADNTRVAATSAAGSALGSVVGTAIGKSIIGTSGATIGAGGAGAA 116

QY 52 -----RVAMAIGGAVLGGIGKIGSQSMDQDKIKLNQSLKVKAG----QVTRWRNP 100

Db 117 AASDRNRTEAAGGA-LGGGAGYTVGNMGNGTNGYTGCAVGAAGGSALGRKVAQDRNY 175

QY 101 DTGNSYSVEPVRTYQRYNKQERR 123
Db 176 DD-----RYDRRYDRDDR 189

RESULT 15
US-08-556-978B-19
; Sequence 19, Application US/08556978B
; Patent No. 6268169
; GENERAL INFORMATION:
; APPLICANT: FARNESTOCK, STEPHEN F.
; TITLE OF INVENTION: NOVEL RECOMBINANTLY PRODUCED
; TITLE OF INVENTION: SPIDER SILK ANALOGS
; NUMBER OF SEQUENCES: 107
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WINDOWS 95
; SOFTWARE: MICROSOFT WORD FOR WINDOWS 95
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/556,978B
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/077,600
; FILING DATE: JUNE 15, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: FLOYD, LINDA AXAMETHY
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: CR-9389-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-892-8112
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 651 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-556-978B-19

Query Match 9.7%; Score 81.5; DB 3; Length 651;
Best Local Similarity 36.8%; Pred. No. 1.2;
Matches 25; Conservative 3; Mismatches 17; Indels 23; Gaps 2;

QY 28 QEVGAATGAVVG-----GVAGQ-----LFGKSGSRVAMAIGGAVLGG 64
Db 491 QGAGAAAVGAGGQGGIRGQGGGQGGYGLGSGSGRGLGGQAGAGAAAAAGGAGQGG 550

QY 65 LIGSKIGQ 72
Db 551 LGGQAGAGQ 558

Search completed: November 5, 2003, 20:16:46
Job time : 11.3523 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 20:15:12 ; Search time 19.0259 Seconds
(without alignments)
1462.395 Million cell updates/sec

Title: US-09-677-374-2

Perfect score: 836

Sequence: 1 MNRGCLQGSSLIISVFLVGG.....IYGTACROPDGRQWVISTEK 162

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 644079 seqs, 171749292 residues

Total number of hits satisfying chosen parameters: 644079

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PTC_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PTC_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09D_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	836	100.0	162	12	US-10-261-446-2
2	815	97.5	161	12	US-10-261-446-4
3	815	97.5	256	12	US-10-261-446-6
4	128.5	15.4	182	15	US-10-127-032-169
5	112	13.4	20	12	US-10-261-446-16
6	83.5	10.0	71	15	US-10-091-724-5
7	81.5	9.7	651	9	US-09-861-597-1
8	81.5	9.7	651	12	US-10-414-760-1
9	80.5	9.6	302	10	US-09-864-864-341
10	79.5	9.5	553	10	US-09-981-353-19
11	79.5	9.5	553	12	US-09-738-630-101
12	79.5	9.5	553	15	US-10-177-233-21
13	79	9.4	618	10	US-09-925-300-1381
14	77.5	9.3	208	12	US-10-029-386-33055
15	77	9.2	489	15	US-10-156-761-8774

16	76.5	9.2	514	9	US-09-815-242-12113	Sequence 12113, A
17	76.5	9.2	890	15	US-10-156-761-14378	Sequence 14378, A
18	76	9.1	346	10	US-09-738-626-5811	Sequence 5811, Ap
19	76	9.1	423	12	US/10/244	Sequence 8, Appli
20	76	9.1	423	15	US/10/013	Sequence 8, Appli
21	76	9.1	423	15	US/10/150	Sequence 8, Appli
22	76	9.1	551	12	US-09-976-782-89	Sequence 89, Appl
23	76	9.1	551	12	US-09-976-782-90	Sequence 90, Appl
24	75	9.0	111	12	US-09-882-237-74	Sequence 74, Appl
25	75	9.0	342	9	US-09-815-242-12104	Sequence 12104, A
26	75	9.0	371	15	US-10-156-761-11322	Sequence 11322, A
27	75	9.0	4588	15	US-10-137-129A-3	Sequence 3, Appli
28	74.5	8.9	274	15	US-10-156-761-9107	Sequence 9107, Ap
29	74.5	8.9	285	10	US-09-978-295A-28	Sequence 28, Appl
30	74.5	8.9	285	10	US-09-978-697-28	Sequence 28, Appl
31	74.5	8.9	285	10	US-09-978-192A-28	Sequence 28, Appl
32	74.5	8.9	285	10	US-09-999-832A-28	Sequence 28, Appl
33	74.5	8.9	285	11	US-09-978-189-28	Sequence 28, Appl
34	74.5	8.9	285	11	US-09-978-608A-28	Sequence 28, Appl
35	74.5	8.9	285	11	US-09-978-585A-28	Sequence 28, Appl
36	74.5	8.9	285	11	US-09-978-191A-28	Sequence 28, Appl
37	74.5	8.9	285	11	US-09-978-403A-28	Sequence 28, Appl
38	74.5	8.9	285	11	US-09-978-564A-28	Sequence 28, Appl
39	74.5	8.9	285	11	US-09-999-833A-28	Sequence 28, Appl
40	74.5	8.9	285	11	US-09-981-915A-28	Sequence 28, Appl
41	74.5	8.9	285	11	US-09-978-824-28	Sequence 28, Appl
42	74.5	8.9	285	11	US-09-918-585A-28	Sequence 28, Appl
43	74.5	8.9	285	11	US-09-978-423A-28	Sequence 28, Appl
44	74.5	8.9	285	11	US-09-978-193A-28	Sequence 28, Appl
45	74.5	8.9	285	11	US-09-999-830A-28	Sequence 28, Appl

ALIGNMENTS

RESULT 1

US-10-261-446-2
; Sequence 2, Application US/10261446
; Publication No. US20030165526A1
; GENERAL INFORMATION:
; APPLICANT: Kuzyk, Michael A.
; APPLICANT: Burian, Jan
; APPLICANT: Kay, William W.
; APPLICANT: Thornton, Julian C.
; TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY IN FISH AGAINST
; TITLE OF INVENTION: RICKETTSIAL DISEASES, AND ASSOCIATED PREVENTATIVE THERAPY
; FILE REFERENCE: 4616-64094
; CURRENT APPLICATION NUMBER: US/10/261,446
; CURRENT FILING DATE: 2002-09-30
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/154,437
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Piscirickettsia salmonis
US-10-261-446-2

Query Match	100.0%;	Score 836;	DB 12;	Length 162;
Best Local Similarity	100.0%;	Pred. No. 1.2e-79;		
Matches 162;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MNRGCLQGSSLIISVFLVGGCAQNFPSRQEVGAATCAVVGAGVAGQVFGKSGRVAMAIGCA	60	
Db	1	MNRGCLQGSSLIISVFLVGGCAQNFPSRQEVGAATCAVVGAGVAGQVFGKSGRVAMAIGCA	60	
QY	61	VLGGLIGSKGSDMDQDKIKLNQSLKVKVAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQ	120	
Db	61	VLGGLIGSKGSDMDQDKIKLNQSLKVKVAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQ	120	

Qy	121	ERRQYCR	FOOKAM	IAGQKE	IYGTAC	RQPDGR	WQVISTEK	162
Db	121	ERRQYCR	FOOKAM	IAGQKE	IYGTAC	RQPDGR	WQVISTEK	162

RESULT 2
US-10-261-446-4
; Sequence 4, Application US/10261446
; Publication No. US20030165526A1
; GENERAL INFORMATION:
; APPLICANT: Kuzyk, Michael A.
; APPLICANT: Burian, Jan
; APPLICANT: Kay, William W.
; APPLICANT: Thornton, Julian C.
; TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY IN FISH AGAINST
; TITLE OF INVENTION: RICKETTSIAL DISEASES, AND ASSOCIATED PREVENTATIVE THERAPY
; FILE REFERENCE: 4616-64094
; CURRENT APPLICATION NUMBER: US/10/261.446
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US 09/677,374
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/154,437
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 161
; TYPE: PRT
; ORGANISM: piscirickettsia salmonis
US-10-261-446-4

	Query Match	97.5%	Score 815;	DB 12;	Length 161;
	Best Local Similarity	98.8;	Pred. No. 1.9e-77;		
	Matches 158;	Conservative 1;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	3	RGCLQGSSLLIIISVFLVGCAGNFSRQFVGATGAVGVGQALFGKSGRVMAGGAVL	62		
Db	2	RGCLQGSSLLIIISVFLVGCAGNFSRQFVGATGAVGVGQALFGKSGRVMAGGAVL	61		
Qy	63	GGLIGSXIQSGMDQODKI KLNQSLSEKVKAGQVTRWRNPDTGNSYSVBPVRYTQRYNKQER	122		
Db	62	GGLIGSXIQSGMDQODKI KLNQSLSEKVKAGQVTRWRNPDTGNSYSVBPVRYTQRYNKQER	121		
Qy	123	ROYCRBFQOKAMTAGQOEIYGTACRQPDGRWQVISTEK	162		
Db	122	ROYCRBFQOKAMTAGQOEIYGTACRQPDGRWQVISTEK	161		

RESULT 3
US-10-261-446-6
; Sequence 6, Application US/10261446
; Publication No. US20030165526A1
; GENERAL INFORMATION:
; APPLICANT: Kuzyk, Michael A.
; APPLICANT: Burian, Jan
; APPLICANT: Kay, William W.
; APPLICANT: Thornton, Julian C.
; TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY IN FISH AGAINST
; TITLE OF INVENTION: RICKETTSIAL DISEASES, AND ASSOCIATED PREVENTATIVE THERAPY
; FILE REFERENCE: 4616-64094
; CURRENT APPLICATION NUMBER: US/10/261.446
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US 09/677,374
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/154,437
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 256
; TYPE: PRT
; ORGANISM: *Piscirickettsia salmonis*
; FEATURE:

```

; NAME/KEY: SIGNAL
; LOCATION: (-95) .. (-1)
US-10-261-446-6

Query Match          97.5%; Score 815; DB 12; Length 256;
Best Local Similarity 98.8%; Pred. No. 3.5e-77;
Matches 158; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 RGCLOGSLIIISVFLVGCNAQNSRQEVGGAATGAVVGGVAGQLFGKSGRVMAIGGAVL 62
    |||||
Db 97 RGCLOGSLIIISVFLVGCNAQNSRQEVGGAATGAVVGGVAGQLFGKSGRVMAIGGAVL 156
    |||||

Qy 63 GGLIGSKI GOSMDQDKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVTRYQYNKQER 122
    |||||
Db 157 GGLIGSKI GOSMDQDKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVTRYQYNKQER 216
    |||||

Qy 123 ROQYCFRFOQKAMIAQKQEIYGTACRQPDGRWQVISTEK 162
    |||||
Db 217 ROQYCFRFOQKAMIAQKQEIYGTACRQPDGRWQVISTEK 256
    |||||

```

```

RESULT 4
US-10-127-032-169
; Sequence 169, Application US/10127032
; Publication No. US20030113742A1
; GENERAL INFORMATION:
; APPLICANT: Whiteley, Marvin
; APPLICANT: Bangera, M. Gita
; APPLICANT: Lory, Stephen
; APPLICANT: Greenberg, Everett Peter
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE MODULATION OF
; TITLE OF INVENTION: BIOFILM FORMATION
; FILE REFERENCE: UIZ-070CP
; CURRENT APPLICATION NUMBER: US/10/127,032
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/285,190
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: US 60/344,142
; PRIOR FILING DATE: 2001-10-24
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 169
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-127-032-169

```

```
Query Match      15.4%; Score 128.5; DB 15; Length 182;
Best Local Similarity 34.8%; Pred. No. 1.5e-05;
Matches 32; Conservative 13; Mismatches 28; Indels 19; Gaps 2;

Qy   28 QEVGAATGAVVGVAQLFGKSGSRVAMAGGAVLGLTSGSKTGQSMDDQDKI----- 80
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db   70 QIAGTAIGAVVGGLLGNQIGGGTGKKIATVAGAVGGYAGNKVQEGNQERDVTYTTETRC 129
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

Qy   81 -KLNQSLKV-----KAGQVTRWRNP 100
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db   130 STVHDSSEKVVGYDVKYMLDGAGQVRMERDP 161
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

RESULT 5
US-10-261-446-16
; Sequence 16, Application US/10261446
; Publication No. US20030165526A1
; GENERAL INFORMATION:
; APPLICANT: Kuzyk, Michael A.
; APPLICANT: Burian, Jan
; APPLICANT: Kay, William W.
; APPLICANT: Thornton, Julian C.
; TITLE OF INVENTION: VACCINES AND AGENTS FOR INDUCING IMMUNITY IN FISH AGAINST
; TITLE OF INVENTION: RICKETTSIAL DISEASES, AND ASSOCIATED PREVENTATIVE THERAPY
; FILE REFERENCE: 4616-64094
; CURRENT APPLICATION NUMBER: US/10/261.446

```

; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: US 09/677,374
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/154,437
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Piscirickettsia salmonis
US-10-261-446-16

```

```

Query Match      13.4%; Score 112; DB 12; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.7e-05;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 110 PVRTYQRYNKQERRQCYCRE 129
Db 1 PVRTYQRYNKQERRQCYCRE 20

```

RESULT 6

```

US-10-091-724-5
; Sequence 5, Application US/10091724
; Publication No. US20030105310A1
; GENERAL INFORMATION:
; APPLICANT: Children's Medical Center Corporation
; APPLICANT: Ashkar, Samy
; TITLE OF INVENTION: Method to Screen Peptide Libraries Using Minicell Display
; FILE REFERENCE: CMCC 820
; CURRENT APPLICATION NUMBER: US/10/091,724
; CURRENT FILING DATE: 2002-03-06
; PRIOR APPLICATION NUMBER: US 60/306,946
; PRIOR FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: 60/274,039
; PRIOR FILING DATE: 2001-03-07
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 71
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: First 71 amino acids of the 17 K antigen of Rickettsia rickettsii
US-10-091-724-5

```

```

Query Match      10.0%; Score 83.5; DB 15; Length 71;
Best Local Similarity 31.3%; Pred. No. 0.23;
Matches 21; Conservative 13; Mismatches 28; Indels 5; Gaps 2;

```

```

QY 9 SLLIIVS---FLVGC--QNFSTRQEVGAATGAVGVAGVAGLFGKSGSRVMAIGGAVLG 63
Db 5 SKIMIIATSLMAACNPGGMNKQGTGTLGGAGGALLGSGFGKGGOLVGVGVGALLG 64
QY 64 GLIGSKI 70
Db 65 AVLGQI 71

```

RESULT 7

```

US-09-861-597-1
; Sequence 1, Application US/09861597
; Patent No. US20020064539A1
; GENERAL INFORMATION:
; APPLICANT: PHILIPPE, Michel
; APPLICANT: GARSON, Jean-Claude
; APPLICANT: ARRAUDEAU, Jean-Pierre
; TITLE OF INVENTION: COSMETIC OR DERMATOLOGICAL COMPOSITION CONTACTING AT
; TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN
; FILE REFERENCE: 6388-0365-0
; CURRENT APPLICATION NUMBER: US/09/861,597

```

```

; CURRENT FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: 09/247,806
; PRIOR FILING DATE: 1999-02-11
; PRIOR APPLICATION NUMBER: FR 98/01614
; PRIOR FILING DATE: 1998-02-11
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Nephila clavipes
US-09-861-597-1

```

```

Query Match      9.7%; Score 81.5; DB 9; Length 651;
Best Local Similarity 36.8%; Pred. No. 6.7;
Matches 25; Conservative 3; Mismatches 17; Indels 23; Gaps 2;

```

```

QY 28 QEVGAATGAVG-----GVAGQ-----LFGKSGSRVMAIGGAVLG 64
Db 491 QGAGAAAAAAGAGQEGIRGQAGGQGGYGLGSGSGRGLGGQAGAAAAAAGGAGQG 550
QY 65 LIGSKIGQ 72
Db 551 LGGQAGQ 558

```

RESULT 8

```

US-10-414-760-1
; Sequence 1, Application US/10414760
; Publication No. US20030192077A1
; GENERAL INFORMATION:
; APPLICANT: Yang, Jianjun G.
; TITLE OF INVENTION: Production of Silk-Like Proteins in Plants
; FILE REFERENCE: BC1014 US NA
; CURRENT APPLICATION NUMBER: US/10/414,760
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: US/09/863,859
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: 60/206968
; PRIOR FILING DATE: 2000-05-25
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 651
; TYPE: PRT
; ORGANISM: Nephila clavipes
US-10-414-760-1

```

```

Query Match      9.7%; Score 81.5; DB 12; Length 651;
Best Local Similarity 36.8%; Pred. No. 6.7;
Matches 25; Conservative 3; Mismatches 17; Indels 23; Gaps 2;

```

```

QY 28 QEVGAATGAVG-----GVAGQ-----LFGKSGSRVMAIGGAVLG 64
Db 491 QGAGAAAAAAGAGQEGIRGQAGGQGGYGLGSGSGRGLGGQAGAAAAAAGGAGQG 550
QY 65 LIGSKIGQ 72
Db 551 LGGQAGQ 558

```

RESULT 9

```

US-09-864-864-341
; Sequence 341, Application US/09864864
; Patent No. US20020102679A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Dillon, David C.
; APPLICANT: Secrist, Heather
; APPLICANT: Lodes, Michael J.
; APPLICANT: Algate, Paul A.

```

APPLICANT: Greenspan, Ralph J.
APPLICANT: Shaw, Paul J.
TITLE OF INVENTION: Methods For Identifying Compounds That

```

; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USING
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33055
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC011553.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.8
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.96
US-10-029-386-33055

Query Match          9.3%; Score 77.5; DB 12; Length 208;
Best Local Similarity 30.8%; Pred. No. 4;
Matches 28; Conservative 11; Mismatches 45; Indels 7; Gaps 3

QY 31 GAATGAVVGGAGQLFGKSGRGVAMAIGGAVLGLLGSKIGOSMDQQDKIKLNQSLEKVK 90
DB 115 GGGTGGGGGTGGGGGTGGGGTGGGGGGGGGGTGVLPGQVPCGGDSTALKQTRPAR 174
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | :
QY 91 AGOVTWRNPDTGNSYSVEPVRTYQRYNKQE 121
DB 175 A-RVTR---PTGRIQ--QPLRSSPHARED 198
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | :

RESULT 15
US-10-156-761-8774
Sequence 8774, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMRRA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 8774
LENGTH: 489
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-8774

Query Match          9.2%; Score 77; DB 15; Length 489;
Best Local Similarity 23.3%; Pred. No. 14;
Matches 31; Conservative 14; Mismatches 42; Indels 46; Gaps 4

QY 29 EVGMATGAVVGGAGQLFGKSGRGVAMAIGGAVLGLLGSKIGOSMDQQDKIKLNQSLEK 88
DB 361 ETRVLTTQAQGHPAPLLFPNGTGRLLAAPDGVLILGTSAGYAGQE-----ET 408
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | :
QY 89 VKAGQV-----TWNRPNDTGNSYS-----VPVRVTYQ----- 115
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | :

```

Db 409 LQAGDLLLLHTDGLVPRWADPRSEAESGGTAAVQRLDLLAPREDGVRTAQHCVRTVV 468
QY 116 -RYNKOERRQOYC 127
Db 469 EEFGETEREDDAC 481

Search completed: November 5, 2003, 20:27:19
Job time : 20.0259 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 20:10:11 ; Search time 11.4715 Seconds
(without alignments)
1358.089 Million cell updates/sec

Title: US-09-677-374-2

Perfect score: 836

Sequence: 1 MNRGCLQGSLIIISVFLVG.....IYGTAQROPDGRWQVISTEK 162

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	296	35.4	159	D33971	rickettsial common
2	293	35.0	159	D33971	Rickettsial common
3	293	35.0	159	D33971	Rickettsial common
4	293	35.0	159	D33971	17K surface antige
5	284.5	34.0	159	C33971	Rickettsial common
6	284	34.0	159	A25972	17K antigen precur
7	128.5	15.4	182	D83169	conserved hypother
8	125.5	15.0	131	A13418	17K surface antige
9	116	13.9	154	D83514	conserved hypother
10	113.5	13.6	155	S23787	outer membrane lip
11	112	13.4	257	B82837	conserved hypother
12	111.5	13.3	142	AD2696	lipA protein (impo
13	111	13.3	125	D97478	lipA protein (impo
14	111	13.3	139	S58234	lipA protein - Rhi
15	110.5	13.2	232	G87629	hypothetical prote
16	108.5	13.0	155	AF0289	probable lipoprote
17	105.5	12.6	155	C64921	outer membrane lip
18	105.5	12.6	155	F90922	probable outer mem
19	105.5	12.6	155	B85771	probable outer mem
20	105.5	12.6	155	A01693	outer membrane lip
21	104	12.4	155	AG0443	outer membrane lip
22	102	12.2	155	I64130	PAL cross-reacting
23	102	12.2	179	AE0644	probable secreted
24	102	12.2	232	AD3350	outer membrane pro
25	101.5	12.1	304	H83636	hypothetical prote
26	99	11.8	179	D85674	hypothetical prote
27	99	11.8	179	H90814	hypothetical prote
28	99	11.8	179	C64855	ycfJ protein - Esc
29	98	11.7	172	G91049	probable outer mem

30	98	11.7	172	2	D85894	probable outer mem
31	98	11.7	172	2	H65026	hypothetical prote
32	98	11.7	179	2	AC0198	probable exported
33	96	11.5	223	2	C82230	probable lipoprote
34	94	11.2	278	2	AB3091	hypothetical prote
35	94	11.2	278	2	H98195	hypothetical prote
36	94	11.2	608	2	T05442	glycine-rich prote
37	93.5	11.2	257	2	F87413	hypothetical prote
38	92	11.0	100	2	G97672	hypothetical prote
39	92	11.0	100	2	AE2897	conserved hypother
40	90	10.8	526	1	KRBOVI	keratin, 54K type
41	90	10.8	1332	2	F69732	PBSX prophage ORF
42	89.5	10.7	220	2	A12910	porin [imported] -
43	89.5	10.7	220	2	G97685	probable outer mem
44	89.5	10.7	691	2	F91251	probable tape meas
45	89	10.6	137	2	B82998	hypothetical prote

ALIGNMENTS

RESULT 1

D33971
rickettsial common antigen precursor - Rickettsia prowazekii
N:Alternate names: 17kD surface antigen; outer membrane protein (omp); RP833
C:Species: Rickettsia prowazekii
C>Date: 16-Mar-1990 #sequence_revision 16-Mar-1990 #text_change 03-Nov-2000
C:Accession: D33971, B71645
R:Anderson, B.E.; Izianabos, T.
J. Bacteriol. 171, 5199-5201, 1989
A:Title: Comparative sequence analysis of a genus-common rickettsial antigen gene.
A:Reference number: A33971; MUID:89359171; PMID:2768201
A:Accession: D33971
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-159 <AND>
A:Cross-references: GB:M28482; NID:g152461
R>Note: the sequence in GenBank entry RIRANT17KC, release 109.0, (PID:g152462) omits the
R:Anderson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alemark, L.
Nature 396, 133-140, 1998
A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A:Reference number: A71630; MUID:99039499; PMID:9823893
A:Accession: B71645
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-159 <AN2>
A:Cross-references: GB:AJ235273; GB:AJ235269; NID:g3861237; PIDN:CAA15258.1; PID:g3861355
R:Experimental source: strain Madrid E
C:Genetics:
A:Gene: omp; RP833
C:Superfamily: rickettsial common antigen
C:Keywords: surface antigen

Query Match 35.4%; Score 296; DB 2; Length 159;

Best Local Similarity 38.1%; Pred. No. 6.6e-19;

Matches 61; Conservative 30; Mismatches 55; Indels 14; Gaps 4;

QY 9 SLLIIISV---FLVGC--AQNFSRQEVGAATGAVGVAGQLFGKSGRVAMAIGAVLG 63

DB 5 SKIMIIAALAAAMLQACNGSGMNMKQGTGTLGGAGGALLGSGFGQKGQLVGVGVALLG 64

QY 64 GIIGSKIGOSMDQDK---IKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNK 119

DB 65 AVLGGQIGASMDQDRRLLELTLSRALESAPSGSNIENWRPNNGHYTPNKTY----- 119

QY 120 QERRQYCYRFFQCKAMIAQKQKQIYGTACRQPDGRWQVIS 159

DB 120 RNSAGYCYRFTVIIGSKQKQKTYGNACRQPDGQWQVNV 159

RESULT 2

B33971
Rickettsial common antigen precursor - Rickettsia conorii

C:Species: Rickettsia conorii
C>Date: 16-Mar-1990 #sequence_revision 16-Mar-1990 #text_change 20-Aug-1999
C:Accession: B33971
R:Anderson, B.E.; Tzianabos, T.
J. Bacteriol. 171, 5199-5201, 1989
A:Title: Comparative sequence analysis of a genus-common rickettsial antigen gene.
A:Reference number: A33971, MUID:89359171, PMID:2768201
A:Accession: B33971
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-159 <AND>
A:Cross-references: GB:M28479; NID:g152463; PIDN:AAA26379.1; PID:g152464
C:Superfamily: rickettsial common antigen

Query Match 35.0%; Score 293; DB 2; Length 159;
Best Local Similarity 38.1%; Pred. No. 1.2e-18;
Matches 61; Conservative 29; Mismatches 56; Indels 14; Gaps 4;

QY 9 SSLIIISV---FLVGC--AQNFSRQEVGAATGAVVGGVAGQLFGKSGSRVMAIGGAVLG 63
DB 5 SKIMIIALATSMLOACNPGGMNKQGTGTLGGAGGALLGSQFGKGGQLVGVGVGALLG 64

QY 64 GLIGSKIGQSMDOODK-----IKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNK 119
DB 65 AVLGGQIGAGMDQDRRLAELTSQRALETAPSGSNVWERNPDNGNGYVTPNKTY----- 119

QY 120 QERRQOYCFEQKAMIAQKQEIYGTACRQPDGRWQVIS 159
DB 120 RNSTGQYCREYTTQTVVIGGKQKAYGNACRQPDGQWQVNV 159

RESULT 3
A33971
Rickettsial common antigen precursor - Rickettsia rickettsii
C:Species: Rickettsia rickettsii
C>Date: 16-Mar-1990 #sequence_revision 16-Mar-1990 #text_change 20-Aug-1999
C:Accession: A33971
R:Anderson, B.E.; Tzianabos, T.
J. Bacteriol. 171, 5199-5201, 1989
A:Title: Comparative sequence analysis of a genus-common rickettsial antigen gene.
A:Reference number: A33971; MUID:89359171; PMID:2768201
A:Accession: A33971
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-159 <AND>
A:Cross-references: GB:M28480; NID:g152457; PIDN:AAA26376.1; PID:g152458
C:Superfamily: rickettsial common antigen

Query Match 35.0%; Score 293; DB 2; Length 159;
Best Local Similarity 38.1%; Pred. No. 1.2e-18;
Matches 61; Conservative 29; Mismatches 56; Indels 14; Gaps 4;

QY 9 SSLIIISV---FLVGC--AQNFSRQEVGAATGAVVGGVAGQLFGKSGSRVMAIGGAVLG 63
DB 5 SKIMIIALATSMLOACNPGGMNKQGTGTLGGAGGALLGSQFGKGGQLVGVGVGALLG 64

QY 64 GLIGSKIGQSMDOODK-----IKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNK 119
DB 65 AVLGGQIGAGMDQDRRLAELTSQRALETAPSGSNVWERNPDNGNGYVTPNKTY----- 119

QY 120 QERRQOYCFEQKAMIAQKQEIYGTACRQPDGRWQVIS 159
DB 120 RNSTGQYCREYTTQTVVIGGKQKAYGNACRQPDGQWQVNV 159

RESULT 4
G97860
17K surface antigen precursor [imported] - Rickettsia conorii (strain Malish 7)
C:Species: Rickettsia conorii
C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 22-Oct-2001
C:Accession: G97860
R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Ro-

A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A:Reference number: A97700; MUID:21442074; PMID:11557893
A:Accession: G97860
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-159 <KUR>
A:Cross-references: GB:AB006914; PIDN:AAL03825.1; PID:g15620425; GSPDB:GN00173
C:Genetics:
A:Gene: omp
C:Superfamily: rickettsial common antigen

Query Match 35.0%; Score 293; DB 2; Length 159;
Best Local Similarity 38.1%; Pred. No. 1.2e-18;
Matches 61; Conservative 29; Mismatches 56; Indels 14; Gaps 4;

QY 9 SSLIIISV---FLVGC--AQNFSRQEVGAATGAVVGGVAGQLFGKSGSRVMAIGGAVLG 63
DB 5 SKIMIIALATSMLOACNPGGMNKQGTGTLGGAGGALLGSQFGKGGQLVGVGVGALLG 64

QY 64 GLIGSKIGQSMDOODK-----IKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNK 119
DB 65 AVLGGQIGAGMDQDRRLAELTSQRALETAPSGSNVWERNPDNGNGYVTPNKTY----- 119

QY 120 QERRQOYCFEQKAMIAQKQEIYGTACRQPDGRWQVIS 159
DB 120 RNSTGQYCREYTTQTVVIGGKQKAYGNACRQPDGQWQVNV 159

RESULT 5
C33971
Rickettsial common antigen precursor - Rickettsia typhi
C:Species: Rickettsia typhi
C>Date: 16-Mar-1990 #sequence_revision 16-Mar-1990 #text_change 20-Aug-1999
C:Accession: C33971
R:Anderson, B.E.; Tzianabos, T.
J. Bacteriol. 171, 5199-5201, 1989
A:Title: Comparative sequence analysis of a genus-common rickettsial antigen gene.
A:Reference number: A33971; MUID:89359171; PMID:2768201
A:Accession: C33971
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-159 <AND>
A:Cross-references: GB:M28481; NID:g152459; PIDN:AAA26377.1; PID:g152460
C:Superfamily: rickettsial common antigen

Query Match 34.0%; Score 284.5; DB 2; Length 159;
Best Local Similarity 39.9%; Pred. No. 6.8e-18;
Matches 55; Conservative 25; Mismatches 49; Indels 9; Gaps 2;

QY 26 SRQEVGAATGAVVGGVAGQLFGKSGSRVMAIGGAVLGLIGSKITGQSNWDOODK----IK 81
DB 27 NKQGTGTLGGAGGALLGSQFGHKGQLVGVGVGALLGAVLGGQIGASLDEQDKLELT 86

QY 82 LNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQERRQOYCFEQKAMIAQKQ 141
DB 87 SQRALESAPSGSNIEWRNPDTGNSHGYVTPNKTY-----RNSTGQYCREYTTQTVVIGGKQ 141

QY 142 EIYGTACRQPDGRWQVIS 159
DB 142 TTYGNACRQPDGQWQVNV 159

RESULT 6
A25972
17K antigen precursor - Rickettsia rickettsii
C:Species: Rickettsia rickettsii
C>Date: 21-May-1988 #sequence_revision 21-May-1988 #text_change 20-Aug-1999
C:Accession: A25972
R:Anderson, B.E.; Regnery, R.L.; Carlone, G.M.; Tzianabos, T.; McDade, J.E.; Fu, Z.Y.;
J. Bacteriol. 169, 2385-2390, 1987
A:Title: Sequence analysis of the 17-kilodalton-antigen gene from Rickettsia rickettsii
A:Reference number: A25972; MUID:87222152; PMID:3108232
A:Accession: A25972

RESULT 8
AI3418
17K surface antigen precursor [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C:Accession: AI3418
R:DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
M.; Mazur, M.; Gotsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letesha,
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:11756688
A:Accession: AI3418
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-131 <KUR>
A:Cross-references: GB:AE008917; PTDN:AA52516.1; PID:cl17983328; GSPDB:GN00190

Db 121 STRAYVQVQGGQIFR 136

RESULT 10

S23787

outer membrane lipoprotein precursor - Yersinia enterocolitica

C:Species: Yersinia enterocolitica

C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 29-Sep-1999

C:Accession: S23787

R:Baeumler, A.J.; Hantke, K.

J. Bacteriol. 174, 1029-1035, 1992

A:Title: A lipoprotein of Yersinia enterocolitica facilitates ferrioxamine uptake in Es

A:Reference number: S23786; MUID:92121089; PMID:1732192

A:Accession: S23787

A:Status: Preliminary

A:Molecule type: DNA
A:Residues: 1-155 <BAE>
A:Cross-references: EMBL:X60448; NID:g48577; PIDN:CAA42977.1; PID:g48579
C:Superfamily: PAL cross-reacting lipoprotein

Query Match 13.6%; Score 113.5; DB 2; Length 155;
Best Local Similarity 24.7%; Pred. No. 0.007;
Matches 37; Conservative 19; Mismatches 43; Indels 51; Gaps 4;

QY 10 SLIIISVFLVGCNQ-----FSRQE-----VGAAT 34
DB 7 AVAAIAAVTGTGCANNNTLSDGFESQAQKQVQTVYTTLLSVRPVTVIQGDDNNVMAIG 66
QY 35 GAVVGAGVQGLFGKSGRVAMATGGAVLGLGSKIGSQNDQDKIKL----- 82
DB 67 GAVLGGFLNTVGGTGRSLATAGAVAGMAGQGVQGANRTDGVQLEVRKDDGTTILV 126

QY 83 --NQSLKVKAGQVTRWRNPDTGNSYSVEP 110
DB 127 VQKQGPTRFSVGQ--RVMLASSGSTVTVSP 154

RESULT 11
B82837
conserved hypothetical protein XF0178 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: B82837
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: B82837
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-257 <SIM>
A:Cross-references: GB:AE003872; GB:AE003849; NID:g9104975; PIDN:AAF82991.1; GSPDB:GN001
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier, H
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranae, E.E.; Laig
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A:Authors: Martins, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, E.C.; Miyaki, C.Y.
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Teuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF0178

Query Match 13.4%; Score 112; DB 2; Length 257;
Best Local Similarity 50.8%; Pred. No. 0.016;
Matches 20; Conservative 8; Mismatches 12; Indels 0; Gaps 0;

QY 31 GAATGAVVGAGVQGLFGKSGRVAMATGGAVLGLGSKI 70
DB 105 GTAIGALIGLVGNQFHGNRKTALTAAGAVAGGFIGNEV 144

RESULT 12
AD2696
lipA protein [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C:Accession: AD2696
R:Wood, D.W.; Secubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Gunthner, D.; Kuttyavin, T.; Levy, R.; Li, M.; McClell

; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AD2696
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-142 <KUR>
A:Cross-references: GB:AE006888; PIDN:AAL41986.1; PID:g17739358; GSPDB:GN00186
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: lipA
A:Map position: circular chromosome

Query Match 13.3%; Score 111.5; DB 2; Length 142;
Best Local Similarity 27.2%; Pred. No. 0.0095;
Matches 41; Conservative 21; Mismatches 60; Indels 29; Gaps 9;

QY 10 SLIIISVFLVGCNQFSRQEVGAATGAVVGAGVQGLFGKSGRVAMATGGAVLGLGSK 69
DB 12 SILCVSM-LSAC-----TTGTRPAG--GSLFGR-SAQPSTPFLANLQGGIVG-K 56

QY 70 IGQSNMDQDKIKL----NQSLKVKAGQVTRWRNPDT-GNSYSVEPVRTYQRYNKQERQ 124
DB 57 SGVELDRGDQTKALEAEYKALETAPVGTPTVWTGDDVKGVVANAP---YQVGN----- 107

QY 125 QYCRFQOKAMTAGQKQEIYFACRQPDGRW 155
DB 108 QNCROYSHTLTVDGRDTRVRGAACRNDGDSW 138

RESULT 13
D97478
lipA protein [imported] - Agrobacterium tumefaciens (strain C58, Cerson)
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
C:Accession: D97478
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorllo, B.; Goldman
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tu
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: D97478
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-125 <KUR>
A:Cross-references: GB:AE007869; PIDN:AAK86781.1; PID:g15155981; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C_1782
A:Map position: circular chromosome

Query Match 13.3%; Score 111; DB 2; Length 125;
Best Local Similarity 27.5%; Pred. No. 0.0092;
Matches 36; Conservative 18; Mismatches 61; Indels 16; Gaps 6;

QY 30 VCAATGAVVGAGVQGLFGKSGRVAMATGGAVLGLGSKIGSQNDQDKIKL----NQS 85
DB 2 LSACTTTGTRPAGGSLFGR-SAQPSTPFLANLQGGIVG-KSGVELDRGDQTKALEAEYKA 59

QY 86 LEKVKAGQVTRWRNPDT-GNSYSVEPVRTYQRYNKQERQYCRFQOKAMTAGQKQEIY 144
DB 60 LETAPVGTPTVWTGDDVKGVVANAP---YQVGN-----QNCROYSHTLTVDGRDTRVR 110

QY 145 GTACRQPDGRW 155
DB 111 GAACRNDGDSW 121

RESULT 14
S58234
lipA protein - Rhizobium leguminosarum

Search completed: November 5, 2003, 20:15:56
Job time : 12.4715 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 20:06:36 ; Search time 6.43523 Seconds
(without alignments)
1183.846 Million cell updates/sec

Title: US-09-677-374-2

Perfect score: 836

Sequence: 1 MNRGCLQGSSLIISVFLVG.....IYGTACRQPDGRWQVISTEK 162

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	296	35.4	159	17KD_RICPR	P16624 rickettsia
2	294	35.2	159	17KD_RICJA	Q52764 rickettsia
3	293	35.0	159	17KD_RICCN	P05372 rickettsia
4	284.5	34.0	159	17KD_RICTY	P22882 rickettsia
5	276	33.0	154	17KD_RICAU	P50928 rickettsia
6	263	31.5	154	17KD_RICAM	P50927 rickettsia
7	261	31.2	154	17KD_RICPA	P50930 rickettsia
8	261	31.2	154	17KD_RICRH	P50931 rickettsia
9	258	30.9	154	17KD_RICMO	P50929 rickettsia
10	157	18.8	80	17KD_RICCA	P29697 rickettsia
11	113.5	13.6	155	PCP_YEREN	P31484 versinia en
12	105.5	12.6	155	SLVE_ECOLI	P55741 escherichia
13	105.5	12.6	155	SLVE_SALTY	Q53549 salmonella
14	102	12.2	155	PCP_HABIN	P10325 haemophilus
15	99	11.8	179	YFCF_ECOLI	P37796 escherichia
16	98	11.7	172	YFGH_ECOLI	P76572 escherichia
17	90	10.8	526	K1CJ_BOVIN	P06394 bos taurus
18	90	10.8	1332	XKDO_BACSU	P45334 bacillus su
19	89	10.6	1585	YQBO_BACSU	P45331 bacillus su
20	87.5	10.5	72	OSMB_SALTY	P37723 salmonella
21	87.5	10.5	783	YKR2_CABEL	P34308 caenorhabdi
22	86.5	10.3	72	OSMB_ECOLI	P17873 escherichia
23	84.5	10.1	243	CYSH_SALTY	P17853 salmonella
24	84.5	10.1	301	STXC_RAT	Q92158 rattus norv
25	84.5	10.1	431	KRE2_CANAL	Q00310 candida alb
26	84.5	10.1	526	VP5_BTW11	P33476 bluetongue
27	83.5	10.0	243	CYSH_SALTY	Q82460 salmonella
28	82	9.8	132	Y615_AQUAE	O66867 aquifex aeo
29	82	9.8	263	CANS_BOVIN	P13135 bos taurus
30	81.5	9.7	541	NUS7_YEAST	P48837 saccharomyc
31	81.5	9.7	747	SPDI_NEFCL	P19837 nephila cla
32	80.5	9.6	467	HEMI_MYCLE	P46724 mycobacteri
33	79.5	9.5	266	CANS_RABIT	P06813 oryctolagus

34	79.5	9.5	359	1	ATPA_BOVIN	P19482 bos taurus
35	79.5	9.5	543	1	ATPA_RAT	P15999 rattus norv
36	79.5	9.5	553	1	ATPA_HUMAN	P25705 homo sapien
37	79	9.4	593	1	K1CJ_HUMAN	P13645 homo sapien
38	78.5	9.4	553	1	ATPO_BOVIN	P19483 bos taurus
39	78.5	9.4	553	1	ATPA_MOUSE	Q03265 mus musculu
40	78	9.3	865	1	VGLB_HSVMD	P18538 marek's dis
41	77.5	9.3	219	1	YIAD_ECOLI	P37665 escherichia
42	77	9.2	269	1	CANS_MOUSE	O88456 mus musculu
43	77	9.2	393	1	CSP_FLABR	P14593 plasmodium
44	77	9.2	429	1	CSP_PLAMA	P13815 plasmodium
45	77	9.2	506	1	ATPA_ANASP	P12405 anabaena ep

ALIGNMENTS

RESULT 1
17KD_RICPR
ID 17KD_RICPR STANDARD; PRT; 159 AA.
AC P16624;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 17 kDa surface antigen precursor.
GN OMP OR RP833.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Madrid E;
RX MEDLINE=89359171; PubMed=2768201;
RA Anderson B.E., Izianabos T.;
RT "Comparative sequence analysis of a genus-common rickettsial antigen
Gene.";
RL J. Bacteriol. 171:5199-5201(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Madrid E;
RX MEDLINE=99039499; PubMed=9823893;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
Sicheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naesslund A.K.,
Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
mitochondria.";
RL Nature 396:133-140(1998).
CC -!- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
anchor (Probable).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; M28482; AAA26378.1; ALT_SEQ.
DR EMBL; AJ235273; CAA15258.1; --
DR PIR; D33971; D33971.
KW Outer membrane; Lipoprotein; Antigen; Signal; Complete proteome.
FT SIGNAL 1 19
FT CHAIN 20 159 17 kDa SURFACE ANTIGEN.
FT LIPID 20 20 N-ACYL DIGLYCERIDE (PROBABLE).
SQ SEQUENCE 159 AA; 16672 MW; A33D404B65EEB071 CRC64;
Query Match 35.4%; Score 296; DB 1; Length 159;
Best Local Similarity 38.1%; Pred. No. 1.5e-18;
Matches 61; Conservative 30; Mismatches 55; Indels 14; Gaps 4;

[illegible]

Db 120 RNSTGQYCREYQTQTVVIGGKQKQAYGNACLPDQG 154

RESULT 9

17KD_RICMO STANDARD; PRT; 154 AA.
AC P50929;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE 17 kDa surface antigen precursor (Fragment).
GN OMP.
OS Rickettsia montana.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsieae; Rickettsia.
OX NCBI_TaxID=33991;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ohio 83-441;
RA Stothard D.R., Ralph D.A., Clark J.B., Fuerst P.A., Pretzman C.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DDAJ databases.
CC -!- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
CC anchor (Probable).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U1017; AAB07705.1;
CC PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
CC KW Outer membrane; Lipoprotein; Antigen; Signal.
CC FT SIGNAL 1 19 BY SIMILARITY.
CC FT CHAIN 20 >154 17 kDa SURFACE ANTIGEN.
CC FT LIPID 20 20 N-ACYL DIGLYCERIDE (PROBABLE).
CC FT NON_TER 154 154
CC SEQUENCE 154 AA; 15881 MW; A09C53B8769E31DA CRC64;

Query Match 30.9%; Score 258; DB 1; Length 154;
Best Local Similarity 36.1%; Pred. No. 2.5e-15;
Matches 56; Conservative 28; Mismatches 57; Indels 14; Gaps 4;
QY 9 SLLIIISV---FLVGC--AQNFSRQEVGAATGAVGVAGQLFGKSGRVAIGAVALG 63
DB 5 SKIMIIAASMLQACNCPGGMKQGTGLAGGAGGALLGSGQGGKQLGVGVGVALG 64
QY 64 GLIGSGIGSMQDQDK----IKLNSLEKVKAGQVTRWNPDTGNSYSVEPVRTYQRYNK 119
DB 65 AVLGGQIGMDEQDRRLAELTSQRALETAPGSGVWNPDPNGYGVTPNKTY----- 119

RESULT 10

17KD_RICCA STANDARD; PRT; 80 AA.
AC P29697;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE 17 kDa surface antigen (Fragment).
GN OMP.
OS Rickettsia canada.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsieae; Rickettsia.
OX NCBI_TaxID=788;
RN [1]
RP SEQUENCE FROM N.A.

QY 120 QRRQOYCHREFQAKMIAGQKQEIYGTACRQPDGR 154
DB 120 RNSTGQYCREYQTQTVVIGGKQKQAYGNACLPDQG 154

RX MEDLINE=92108069; PubMed=1729713;
RA Azad A.F., Sacci J.B. Jr., Nelson W.M., Dasch G.A.,
RA Schmidtman E.T., Carl M.;
RT "Genetic characterization and transovarial transmission of a
RT typhus-like rickettsia found in cat fleas";
RL Proc. Natl. Acad. Sci. U.S.A. 89:43-46(1992).
CC -!- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
CC anchor (Probable).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M82879; -; NOT ANNOTATED CDS.
CC PROSITE; PS000437; PROK_LIPOPROT.
CC DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; PARTIAL.
CC KW Outer membrane; Lipoprotein; Antigen.
CC FT NON_TER 1 1
CC FT NON_TER 80 80
CC SEQUENCE 80 AA; 8372 MW; AD289A48EAB19E0E CRC64;
QY 43 GOLFGKSGRVAIGAVALGGLIGSKTGQSDQDQK----IKLNSLEKVKAGQVTRWR 98
DB 1 GSQFGKQGLQIGVAGALLGAILGNQIGAGMDEQDRRLAELTSQRALETTPSGTSIEWR 60
QY 99 NPDTCNSYSVEPVRTYQ 115
DB 61 NPDNGNGYGVTPSKTYK 77
RESULT 11
PCP_YEREN
ID PCP_YEREN STANDARD; PRT; 155 AA.
AC P31484;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Outer membrane lipoprotein pcp precursor.
GN PCP OR PCPY.
OC Yersinia enterocolitica.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=630;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 51872 / WA-C / Serotype O:8;
RX MEDLINE=92121089; PubMed=1732192;
RA Baumler A.J., Hantke K.;
RT "A lipoprotein of Yersinia enterocolitica facilitates ferrioxamine
RT uptake in Escherichia coli";
RL J. Bacteriol. 174:1029-1035(1992).
CC -!- FUNCTION: FACILITATES FERRIOXAMINE UPTAKE.
CC -!- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
CC anchor.
CC -!- SIMILARITY: TO E. COLI AND S. TYPHIMURIUM SLVB AND TO
CC H. INFLUENZAE PCP.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

```

DR EMBL; X60448; CAB42977.1; -.
DR PIR; S23787; S23787.
DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.
KW Outer membrane; Lipoprotein; Signal.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 155 OUTER MEMBRANE LIPOPROTEIN PCP.
FT LIPID 18 18 N-ACYL DIGLYCERIDE (POTENTIAL).
SQ SEQUENCE 155 AA; 15362 MW; 8AD6BE2132E849FA CRC64;

Query Match 13.6%; Score 113.5; DB 1; Length 155;
Best Local Similarity 24.7%; Pred. No. 0.0053;
Matches 37; Conservative 19; Mismatches 43; Indels 51; Gaps 4;

QY 10 SLIIISVFLVGCAN-----FSRQE-----VGRAT 34
DB 7 AVAIAAIVLTGCGANNNTLSGDVFSQAQKQVQTVTGTLLSVRPTVIOGGDDNNVWGAG 66
QY 35 GAVGVGVAGOLFCKGSRVMAIGGAVLGLIGSKIGSQMDQDKL-----82
DB 67 GAVLGGFLGTVGGTGRSLATAAGAVAGMGAGQGVQGMNRTDGVQLEVRKDDGTTILV 126
QY 83 --NQSLKVKAGQVTRWRNPDTGNSYSVSP 110
DB 127 VQKQGPTRFSVGO--RWLASSGSTVTVP 154

RESULT 12
SLYB_ECOLI
ID SLYB_ECOLI STANDARD; PRT; 155 AA.
AC P55741; P76183;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Outer membrane lipoprotein slyB precursor.
GN SLYB OR B1641 OR Z2655 OR ECS2350 OR SF1668.
OS Escherichia coli,
OS Shigella flexneri
OS Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
ON NCBI_TaxID=562, 83334, 623;
[1]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli;
RX MEDLINE=96133688; PubMed=8544813;
RA Ludwig A., Tengel C., Bauer S., Bubert A., Benz R., Mollenkopf H.-J.,
RA Goebel W.;
RT "slyA, a regulatory protein from Salmonella typhimurium, induces a
RL haemolytic and pore-forming protein in Escherichia coli.";
RN Mol. Gen. Genet. 249:474-486(1995).
[2]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
[3]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12;
RX MEDLINE=97251357; PubMed=9097039;
RA Aiba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,
RA Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T.,
RA Motomura K., Nakade S., Nakamura Y., Nashimoto H., Nishio Y.,
RA Oshima T., Saito N., Sampaio G., Seki Y., Sivasubraman S.,
RA Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C.,
RA Yamamoto Y., Horiuchi T.;
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RL corresponding to the 28.0-40.1 min region on the linkage map.";
RN

DNA Res. 3:363-377(1996).
[4]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Weich R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
[5]
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156233; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,
RA Kubara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
[6]
RP SEQUENCE FROM N.A.
RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang X., Shen Y., Lu W., Wang J., Liu H.,
RA Sun L., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).
CC -! SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
CC anchor (Potential).
CC -! SIMILARITY: TO S.TYPHIMURIUM SLYB, H.INFLUENZAE PCP AND
CC Y. ENTEROCOLITICA PCP.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000259; AAC74713.1; -.
CC EMBL; D90807; BAA15402.1; -.
CC EMBL; AE005387; AAG56630.1; -.
CC EMBL; AF002558; BAB35773.1; -.
CC EMBL; AE015187; AAN43250.1; -.
CC PIR; B85771; B85771.
CC PIR; C64921; C64921.
CC PIR; F90922; F90922.
CC EcoGene; EG13409; slyB.
CC PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.
KW Outer membrane; Lipoprotein; Signal; Complete proteome.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 155 OUTER MEMBRANE LIPOPROTEIN SLYB.
FT LIPID 18 18 N-ACYL DIGLYCERIDE.
FT CONFLICT 98 98 A -> T (IN REF. 2).
SQ SEQUENCE 155 AA; 15602 MW; 543EB8A4069A5FA3 CRC64;

Query Match 12.6%; Score 105.5; DB 1; Length 155;
Best Local Similarity 22.8%; Pred. No. 0.025;
Matches 33; Conservative 21; Mismatches 42; Indels 49; Gaps 4;

QY 13 IISVFLVGCAN-----FSRQE-----VGRAT 37
DB 10 MVGLSLVGCNVNDTSLSGDVITASEAKQVQNSYGTIVNRPVIOGGDDSNVIGAGV 69

```

QY 38 VGVAGQLFGKSGRVAMAIGGAVLGLGKIGKIGQSMDDQDKIKLQSLKVKAGQVTRW 97
 Db 70 LGGFLGNTVGGTGRSLATAAGAVAGVAGQGVQSAAMNTQGVLE--EIRK----- 118
 QY 98 RNPDTGNSYSVEPRTYQRYNKQER 122
 Db 119 ---DDGNTIMVVKQGNTRFSPQOR 140

RESULT 13
 ID SLVB_SALTY STANDARD; PRT; 155 AA.
 AC Q53549;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Outer membrane lipoprotein slyb precursor.
 GN SLVB OR STW1445 OR STY1677 OR T1313.
 OS Salmonella typhimurium, and
 OS Salmonella typhi.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=602, 601;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.typhimurium;
 RX MEDLINE=96133688; PubMed=8544813;
 RA Ludwig A., Tengel C., Bauer S., Buber A., Benz R., Mollenkopf H.-J.,
 RA Goebel W.;
 RT "SlyA, a regulatory protein from Salmonella typhimurium, induces a
 RT haemolytic and pore-forming protein in *Escherichia coli*.";
 RL Mol. Gen. Genet. 249:474-486(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
 RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium
 RT LT2.";
 RL Nature 413:852-856(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.typhi; STRAIN=CT18;
 RX MEDLINE=21534947; PubMed=11677608;
 RA Parkhill J., Dougan G., James K.D., Holden M.T.G., Sebahia M.,
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
 RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
 RA Whitehead S., Barrett B.G.;
 RT "Complete genome sequence of a multiple drug resistant *Salmonella*
 RT enterica serovar Typhi CT18.";
 RL Nature 413:848-852(2001).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.typhi; STRAIN=Ty2 / ATCC 700931;
 RX MEDLINE=22531367; PubMed=12644504;
 RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
 RA Burland V., Kodoyanni V., Schwartz D.C., Blattner F.R.;
 RT "Comparative genomics of *Salmonella enterica* serovar Typhi strains Ty2
 RT and CT18.";
 RL J. Bacteriol. 185:2330-2337(2003).
 CC -!- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
 CC anchor (potential).
 CC -!- SIMILARITY: TO E.COLI SLVB, H.INFLUENZAE PCP AND Y. ENTEROCOLITICA
 CC PCP.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; S80790; AAB35871.2; -;
 CC EMBL; AE008762; AAL20367.1; -;
 CC EMBL; AL627271; CAD01922.1; -;
 CC EMBL; AE016838; AAO68963.1; -;
 CC STYGENE; PS00013; PROKAR LIPOPROTEIN; 1.
 KW Outer membrane; Lipoprotein; Signal; Complete proteome.
 FT SIGNAL 1 17 POTENTIAL.
 FT CHAIN 18 155 OUTER MEMBRANE LIPOPROTEIN SLVB.
 FT LIPID 18 18 N-ACYL DIGLYCERIDE.
 SQ SEQUENCE 155 AA; 15548 MW; 82FDDCDBABD55A7 CRC64;
 Query Match 12.6%; Score 105.5; DB 1; Length 155;
 Best Local Similarity 24.6%; Pred. No. 0.025; Indels 33; Gaps 3;
 Matches 34; Conservative 19; Mismatches 52;
 QY 4 GCLQGSSL-----IIISVFLVCAQNFQSRQEVGAATGAVVGAGQ 44
 Db 17 GCVNDSLSGDVYTASEAKQVNTYGTIVNRPVQIOGGDSDNVIGAGVILGGFLGN 76
 QY 45 LFGKSGRVAMAIGGAVLGLGKIGKIGQSMDDQDKIKLQSLKVKAGQVTRWPTDGN 104
 Db 77 TIGGCGTGRSLATAAGAVAGVAGQGVQSAAMNTQGVLE--EIRK-----DDGN 122
 QY 105 SYSVEPRTYQRYNKQER 122
 Db 123 TIMVVKQGNTRFSPQOR 140
 RESULT 14
 PCP_HAEIN
 ID PCP_HAEIN STANDARD; PRT; 155 AA.
 AC P10325;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Outer membrane lipoprotein PCP precursor (15 kDa lipoprotein) (PAL
 DE cross-reacting lipoprotein).
 GN PCP OR LPP OR H11579.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae; Haemophilus.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88115138; PubMed=2828309;
 RA Deich R.A., Metcalf B.J., Finn C.W., Farley J.E., Green B.A.;
 RT "Cloning of genes encoding a 15,000-dalton peptidoglycan-associated
 RT outer membrane lipoprotein and an antigenically related 15,000-dalton
 RT protein from *Haemophilus influenzae*.";
 RL J. Bacteriol. 170:489-498(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RD / KW20 / ATCC 51907;
 RX MEDLINE=95350830; PubMed=7542800;
 RA Fieischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McAnney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.P., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.W., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;

RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RD Rd."; Science 269:496-512(1995).
RL -!- SUBCELLULAR LOCATION: Attached to the outer membrane by a lipid
CC anchor.
CC -!- SIMILARITY: TO E. COLI AND S. TYPHIMURIUM SLyb AND TO
CC Y. ENTEROCOLITICA PCP.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M18877; AAA24938.1; -;
DR EMBL; U32832; AAC23228.1; -;
DR PIR; I64130; I64130.
DR TIGR; H11579; -;
DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.
KW Outer membrane; Lipoprotein; Signal; Complete proteome.
FT SIGNAL 1 18
FT CHAIN 19 155 OUTER MEMBRANE LIPOPROTEIN PCP.
FT LIPID 19 15 N-ACYL DIGLYCERIDE.
FT CONFLICT 135 143 CSLVAEFVF -> VAGRRVRI (IN REF. 1).
FT SEQUENCE 155 AA; 15425 MW; D7880327FCF0C985 CRC64;
SQ
Query Match 12.2%; Score 102; DB 1; Length 155;
Best Local Similarity 39.6%; Pred. No. 0.051;
Matches 21; Conservative 9; Mismatches 23; Indels 0; Gaps 0;
QY 30 VGNATGAVGVAGQLFGKSGRVAIAIGGAVLGGIGSKIGSQMDQDKIKL 82
DB 62 VGFLGGALGGIAGTIGGGRGQAIAAVGGAIGGATAGSKIKBKMSQVNGAEL 114
RESULT 15
YCFJ_ECOLI STANDARD; PRT; 179 AA.
ID YCFJ_ECOLI
AC F37796; P75951;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hydrothermal protein ycfJ
GN YCFJ OR B1110 OR C1383 OR SF1114.
OS Escherichia coli,
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562, 217992, 623;
[1]
RN
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick J.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
[2]
RN
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sempel G., Seki Y., Tagami H., Takenmoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome

RT corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
[3]
RN
RP SEQUENCE FROM N.A.
RC SPECIES=E.coli; STRAIN=O6:H1 / CFT073 / ATCC 700928;
RX MEDLINE=22388234; PubMed=12471157;
RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT "Extensive mosaic structure revealed by the complete genome sequence
RT of uropathogenic Escherichia coli.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
[4]
RN
RP SEQUENCE OF 1-63 FROM N.A.
RC SPECIES=E.coli; STRAIN=K12;
RX MEDLINE=81236546; PubMed=6265208;
RA Young J.G., Rogers B.L., Campbell H.D., Jaworski A., Shaw D.C.;
RT "Nucleotide sequence coding for the respiratory NADH dehydrogenase of
RT Escherichia coli. UUG initiation codon.";
RL Eur. J. Biochem. 116:165-170(1981).
[5]
RN
RP SEQUENCE FROM N.A.
RC SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590;
RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).
[6]
RN
RP IDENTIFICATION.
RC SPECIES=E.coli;
RX MEDLINE=95075659; PubMed=7984428;
RA Borodovsky M., Rudd K.E., Koonin E.V.;
RT "Intrinsic and extrinsic approaches for detecting genes in a
RT bacterial genome.";
RL Nucleic Acids Res. 22:4756-4767(1994).
CC -!- SIMILARITY: TO RICKETTSIA 17 KDa SURFACE ANTIGEN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000211; AAC74194.1; -;
DR EMBL; D90746; BAA35925.1; -;
DR EMBL; AE016759; AAN79853.1; ALT INIT.
DR EMBL; V00306; -; NOT ANNOTATED CDS.
DR EMBL; AE015138; AAN42732.1; ALT_INIT.
DR PIR; C64855; C64855.
DR EColGene; EG12444; ycfJ.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 5 25 POTENTIAL.
SQ SEQUENCE 179 AA; 18920 MW; BAE5B0DB56D45609 CRC64;
Query Match 11.8%; Score 99; DB 1; Length 179;
Best Local Similarity 27.1%; Pred. No. 0.11;
Matches 26; Conservative 12; Mismatches 32; Indels 26; Gaps 1;
QY 31 GAATGAVGVAGQLFGKSGRVAIAIGGAVLGGIGSKIGSQMDQDKIKLNSLEKVK 90
DB 73 GSVLGAVGAVGVIGHQFGGGRGKDVATVVGALGGVAGNQTGSLQESD----- 120
QY 91 ACQVTRWRNPDTGNSYSVEPVTYQRYNQRQQY 126
DB 121 -----TYTTTQQRCKTVYDKSKMLGY 142

Search completed: November 5, 2003, 20:13:25
Job time : 7.43523 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2003, 20:09:46 ; Search time 23.7824 Seconds
(without alignments)
1757.794 Million cell updates/sec

Title: US-09-677-374-2

Perfect score: 836

Sequence: 1 MNRGCLQGSSLIISVFLVG.....IVGTACRQPDGRWQVISTEK 162

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	836	100.0	162 2 Q9F9K8	Q9F9K8 pisciricket
2	313.5	37.5	148 2 O54381	O54381 rickettsia
3	291	34.8	159 2 Q9F9F2	Q9F9F2 rickettsia
4	266.5	31.9	137 2 O52252	O52252 rickettsia
5	260.5	31.2	144 2 Q9K2N6	Q9K2N6 male-killin
6	259.5	31.0	137 2 O31065	O31065 rickettsia
7	259.5	31.0	144 2 Q9K4W8	Q9K4W8 male-killin
8	252	30.1	154 2 O53154	O53154 rickettsia
9	247	29.5	151 2 Q9F9Q9	Q9F9Q9 rickettsia
10	242.5	29.0	131 2 O8KLT4	O8KLT4 rickettsia
11	242.5	29.0	131 2 O8KLT5	O8KLT5 rickettsia
12	237.5	28.4	131 2 Q9F0Q1	Q9F0Q1 rickettsia
13	237.5	28.4	131 2 Q52637	Q52637 rickettsia
14	236.5	28.3	131 2 O8KLU2	O8KLU2 rickettsia
15	236.5	28.3	131 2 Q91522	Q91522 rickettsia
16	223.5	26.7	105 2 O31208	O31208 rickettsia

ALIGNMENTS

RESULT 1

Q9F9K8	Q9F9K8	PRELIMINARY;	PRT;	162 AA.
ID	AC	Q9F9K8;		
DT	01-MAR-2001	(TREMBLrel. 16, Created)		
DT	01-MAR-2001	(TREMBLrel. 16, Last sequence update)		
DT	01-MAR-2001	(TREMBLrel. 16, Last annotation update)		
DE	17 Kda antigen.			
GN	OSPA.			
OS	Piscirickettsia salmonis.			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Thiotrichales;			
OC	Piscirickettsiaceae; Piscirickettsia.			
OX	NCBI_TaxID=1238;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAINS=LF-89;			
RA	Kuzyk M.A., Burian J., Thornton J.C., Kay W.W.;			
RT	"Identification of a genus-common Rickettsial surface antigen in the salmonid pathogen Piscirickettsia salmonis."			
RL	Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF184152; AAG17000.1;			
SQ	SEQUENCE 162 AA; 17661 MW; DDE99B6FD94A527E CRC64;			

Query Match	100.0%;	Score 836;	DB 2;	Length 162;
Best Local Similarity	100.0%;	Pred. No. 3.1e-67;		
Matches 162;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1 MNRGCLQGSSLIISVFLVGCAQNFROEVGAATGAVVGGVAGQLFGKSGRVAMAIGGA 60			
Db	1 MNRGCLQGSSLIISVFLVGCAQNFROEVGAATGAVVGGVAGQLFGKSGRVAMAIGGA 60			
QY	61 VLGLIGSKIGOSMDQDKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQ 120			
Db	61 VLGLIGSKIGOSMDQDKIKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQ 120			
QY	121 ERQOYCFEQKAMIAQKQEIYGTACRQPDGRWQVISTEK 162			
Db	121 ERQOYCFEQKAMIAQKQEIYGTACRQPDGRWQVISTEK 162			

Qy 64 GLIGSKTIGQSMDOQDK-----IKLNQSLKVKAGQVTRWRNPDTGNSYVPEVRTYQRYNK 119
 Db 65 AVLGGQICGAGMDEQDQLAEITSQRALEATPSGTSSVWRNPDNGNHGYVTPNKTY----- 119

Qy 120 QERRQOQYCRBFOQKAMTAGQKOEIYGTACRQPDGRWVOVIS 159
 Db 120 RNSTGQYCREYTVTVIGGQKQKAYGNACRQPDGLWQVNV 159

RESULT 4
 O52252 PRELIMINARY; PRT; 137 AA.
 AC O52252;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE 17 kDa antigen (Fragment).
 OS Rickettsia cooleyi.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsiae; Rickettsia.
 OX NCBI_TaxID=69410;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Billings A.N., Teltow G.J., Walker D.H.;
 RT "Molecular characterization of a novel spotted fever group rickettsial
 RT species from Ixodes scapularis in Texas";
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 FT EMBL; AF031534; AAB95267.1; -.
 FT NON TER 1
 SQ SEQUENCE 137 AA; 14215 MW; A27597A9AFD85FC3 CRC64;

Query Match 31.9%; Score 266.5; DB 2; Length 137;
 Best Local Similarity 39.6%; Pred No. 2.6e-16;
 Matches 53; Conservative 23; Mismatches 49; Indels 9; Gaps 2

Qy 23 QNFSRQVGAATGAVVGVGAGQIFGKSGSRVAMAIGGAVLGLIGSKIGQSMDOQDK--- 79
 Db 7 RGMNKGQGTLLGGAGGALLGSGQFGKGGQLGVGVGALLGAVLGGQIGAGMDEQDRRLA 66

Qy 80 -IKLNQSLKVKAGQVTRWRNPDTGNSYVPEVRTYQRYNKQERRQOQYCRBFOQKAMTAG 138
 Db 67 ELTSQRALEAAPSGSSTEWRNPONGNYGYVTPNKTY-----RNSTGQYCREYTVTVIGG 121

Qy 139 QKOEIYGTACRQPD 152
 Db 122 KQKQKAYGNACRQPD 135

RESULT 5
 Q9K2N6 PRELIMINARY; PRT; 144 AA.
 ID Q9K2N6
 AC Q9K2N6;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE 17kDa antigen (17 kDa antigen) (fragment).
 OS male-killing Rickettsia from Adalia bipunctata.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsiae; Rickettsia.
 OX NCBI_TaxID=38028;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Schlenker H.J.G.V.D., Habis M., Sloggett J.J., Webberley M.K.,
 RA Berstrand D., Hurst G.D.D., Majerus M.E.N.;
 RT "On the evolution of male-killing: Monophyletic origin and horizontal
 RT transfer of male-killing Rickettsia (a-Proteobacteria) from two con-
 RT generic ladybirds, Adalia bipunctata L. and A. decempunctata L.
 RT (Coleoptera: Coccinellidae)".
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ269518; CAB96383.1; -.
 DR ENBL; AJ269517; CAB96382.1; -.
 FT NON TER 1
 FT NON TER 144 144

```

SQ SEQUENCE 144 AA; 14785 MW; C8254739CCA56AE7 CRC64;
Query Match
Best Local Similarity 31.2%; Score 260.5; DB 2; Length 144;
Matches 52; Conservative 22; Mismatches 48; Indels 9; Gaps 2;

QY 26 SRQEVGAATGAVGVGAGQLFKGSGRVRMAIGGAVLGLGSKIGQSMDDQDK----IK 81
Db 17 NRQGTGTLGGAGGALLGSQFGKRGQLVGVGVGALLGAVLGGIGAGMDEQDRRLAELT 76

QY 82 LNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKRQYCRFQOKAMTAGOKO 141
Db 77 SORALEAPSGSNVWRNPNGNHGYVTPNKTY-----RNSTGQYCREYTQTVVIGGKQO 131

QY 142 EIYGTACROPD 152
Db 132 KAYGNACROPD 142

RESULT 6
O31065 PRELIMINARY; PRT; 137 AA.
AC O31065; Q9HW02;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE 17 kDa antigen (17 kDa protein) (Fragment).
OS Rickettsia honei.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiidae; Rickettsia.
OX NCBI_TaxID=37816;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TT-118;
RA Billings A.N., Yu X.-J., Teal P.D., Walker D.H.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99045882; PubMed=9828442;
RA Stenos J., Roux V., Walker D., Raoult D.;
RT "Rickettsia honei sp. nov., the aetiological agent of Flinders Island
RT spotted fever in Australia.";
RL Int. J. Syst. Bacteriol. 48:1399-1404 (1998).
DR EMBL; AF027124; AAB81846.1; -.
DR EMBL; AF060706; AAD20231.1; -.
DR EMBL; AF060704; AAD20230.1; -.
FT NON_TER 1
SQ SEQUENCE 137 AA; 14167 MW; 75BC1D0D745B428C CRC64;

Query Match
Best Local Similarity 31.0%; Score 259.5; DB 2; Length 137;
Matches 52; Conservative 22; Mismatches 48; Indels 9; Gaps 2;

QY 26 SRQEVGAATGAVGVGAGQLFKGSGRVRMAIGGAVLGLGSKIGQSMDDQDK----IK 81
Db 10 NRQGTGTLGGAGGALLGSQFGKRGQLVGVGVGALLGAVLGGIGAGMDEQDRRLAELT 69

QY 82 LNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKRQYCRFQOKAMTAGOKO 141
Db 70 SORALEAPSGSNVWRNPNGNHGYVTPNKTY-----RNSTGQYCREYTQTVVIGGKQO 124

QY 142 EIYGTACROPD 152
Db 125 KAYGNACROPD 135

RESULT 7
Q9K4W8 PRELIMINARY; PRT; 144 AA.
AC Q9K4W8;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)

```

```

DE 17 kDa antigen (Fragment).
OS male-killing Rickettsia from Adalia decempunctata.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiidae; Rickettsia.
OX NCBI_TaxID=120393;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20575219; PubMed=11133455;
RA Schlenker H.J.G.V.D., Habis M., Sloggett J.J., Webberley M.K.,
RA Berland D., Hurst G.D., Majerus M.E.N.;
RT "Incidence of male-killing Rickettsia spp. (alpha-Proteobacteria) in
RT the Ten-Spot ladybird Beetle Adalia decempunctata L. (Coleoptera:
RT Coccinellidae).";
RL Appl. Environ. Microbiol. 67:270-277 (2001).
DR EMBL; AJ269516; CAB96381.1; -.
FT NON_TER 1
FT NON_TER 144
SQ SEQUENCE 144 AA; 14801 MW; C825472F16A56AE7 CRC64;

Query Match
Best Local Similarity 31.0%; Score 259.5; DB 2; Length 144;
Matches 52; Conservative 22; Mismatches 48; Indels 9; Gaps 2;

QY 26 SRQEVGAATGAVGVGAGQLFKGSGRVRMAIGGAVLGLGSKIGQSMDDQDK----IK 81
Db 17 NRQGTGTLGGAGGALLGSQFGKRGQLVGVGVGALLGAVLGGIGAGMDEQDRRLAELT 76

QY 82 LNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKRQYCRFQOKAMTAGOKO 141
Db 77 SORALEAPSGSNVWRNPNGNHGYVTPNKTY-----RNSTGQYCREYTQTVVIGGKQO 131

QY 142 EIYGTACROPD 152
Db 132 KSYGNACROPD 142

RESULT 8
Q53154 PRELIMINARY; PRT; 154 AA.
AC Q53154;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE (Clone PRB F15F 1), 5' end CDS (Fragment).
OS Rickettsia sp.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC Rickettsiaceae; Rickettsiidae; Rickettsia.
OX NCBI_TaxID=789;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93084757; PubMed=1452660;
RA Baird R.W., Lloyd M., Stenos J., Ross B.C., Stewart R.S., Dwyer B.;
RT "Characterization and comparison of Australian human spotted fever
RT group rickettsiae.";
RL J. Clin. Microbiol. 30:2896-2902 (1992).
DR EMBL; M99391; AAA73386.1; -.
FT NON_TER 154
SQ SEQUENCE 154 AA; 15849 MW; F5C35855EDB439D2 CRC64;

Query Match
Best Local Similarity 30.1%; Score 252; DB 2; Length 154;
Matches 55; Conservative 29; Mismatches 57; Indels 14; Gaps 4;

QY 9 SLLIISV---FLVGC--AQNPSRQEVGAATGAVGVGAGQLFKGSGRVRMAIGGAVL 63
Db 5 SKIMIATATSMLOACNGPGMKNKQGTGTLGGAGGALLGSQFGKRGQLVGVGVGALLG 64

QY 64 GLIGSKIGQSMDDQDK----IKLNQSLKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYN 119
Db 65 AVLGGIGAGMDEQDRRLAELTSQLAETAPSGSNVWRNPNGNSYGYVTPNKTYRNSTG 124

QY 120 QRRQYCRFQOKAMTAGOKQEIYGTACROPDGR 154

```

Query Match	29.0%	Score 242.5;	DB 2;	Length 131;
Best Local Similarity	38.6%;	Pred. No. 3.5e-14;		
Matches 49;	Conservative 22;	Mismatches 47;	Indels 9;	Gaps 2;

RT "A new SFG rickettsia isolated from fleas.";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]

RC SEQUENCE FROM N.A.
 RC STRAIN=California 2;
 RA Roux V., Raoult D.;

RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF210693; AAG48554.1; -.
 FT NON_TER 1
 FT NON_TER 131
 SQ SEQUENCE 131 AA; 13374 MW; 23C8819B29FF860 CRC64;

Query Match 28.4%; Score 237.5; DB 2; Length 131;
 Best Local Similarity 37.8%; Pred. No. 9.8e-14;
 Matches 48; Conservative 22; Mismatches 48; Indels 9; Gaps 2;

QY 26 SRQEVGAATGAVGVGAGQVTRWRNPDGNSYSVEPVRYQYKQERRQQYCRFQOKAMTAGOKQ 141
 Db 10 NKQGTGTTLLGGAGGALLGSGQFGKQGVGVGALLGAVLGGQIGAGMDEQDRRLAELT 69

QY 82 LNQSLEKVKAGQVTRWRNPDGNSYSVEPVRYQYKQERRQQYCRFQOKAMTAGOKQ 141
 Db 70 SORALEAPSGSVNWRPNNGHGYTPNKTY-----RNSTGQYCREYTVTVVIGGKQ 124

QY 142 EYGTAC 148
 Db 125 KAYGNAC 131

RESULT 13
 Q52637 PRELIMINARY; PRT; 131 AA.

AC Q52637;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE 17 kDa antigen (Fragment).

OS Rickettsia sp.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsiidae; Rickettsia.
 OX NCBI_TaxID=789;

Query Match 28.4%; Score 237.5; DB 2; Length 131;
 Best Local Similarity 37.8%; Pred. No. 9.8e-14;
 Matches 48; Conservative 22; Mismatches 48; Indels 9; Gaps 2;

QY 26 SRQEVGAATGAVGVGAGQVTRWRNPDGNSYSVEPVRYQYKQERRQQYCRFQOKAMTAGOKQ 141
 Db 10 NKQGTGTTLLGGAGGALLGSGQFGKQGVGVGALLGAVLGGQIGAGMDEQDRRLAELT 69

QY 82 LNQSLEKVKAGQVTRWRNPDGNSYSVEPVRYQYKQERRQQYCRFQOKAMTAGOKQ 141
 Db 70 SORALEAPSGSVNWRPNNGHGYTPNKTY-----RNSTGQYCREYTVTVVIGGKQ 124

QY 142 EYGTAC 148
 Db 125 KAYGNAC 131

RESULT 13
 Q52637 PRELIMINARY; PRT; 131 AA.

AC Q52637;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE 17 kDa antigen (Fragment).

OS Rickettsia sp.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsiidae; Rickettsia.
 OX NCBI_TaxID=789;

Query Match 28.4%; Score 237.5; DB 2; Length 131;
 Best Local Similarity 37.8%; Pred. No. 9.8e-14;
 Matches 48; Conservative 22; Mismatches 48; Indels 9; Gaps 2;

QY 26 SRQEVGAATGAVGVGAGQVTRWRNPDGNSYSVEPVRYQYKQERRQQYCRFQOKAMTAGOKQ 141
 Db 10 NKQGTGTTLLGGAGGALLGSGQFGKQGVGVGALLGAVLGGQIGAGMDEQDRRLAELT 69

QY 82 LNQSLEKVKAGQVTRWRNPDGNSYSVEPVRYQYKQERRQQYCRFQOKAMTAGOKQ 141
 Db 70 SORALEAPSGSVNWRPNNGHGYTPNKTY-----RNSTGQYCREYTVTVVIGGKQ 124

QY 142 EYGTAC 148
 Db 125 KAYGNAC 131

RESULT 14
 Q8KLU2 PRELIMINARY; PRT; 131 AA.

AC Q8KLU2;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Outer membrane protein (Fragment).
 GN 17 kDa.
 OS Rickettsia helvetica.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsiidae; Rickettsia.
 OX NCBI_TaxID=35789;

Query Match 28.3%; Score 236.5; DB 2; Length 131;
 Best Local Similarity 37.8%; Pred. No. 1.2e-13;
 Matches 48; Conservative 22; Mismatches 48; Indels 9; Gaps 2;

QY 26 SRQEVGAATGAVGVGAGQVTRWRNPDGNSYSVEPVRYQYKQERRQQYCRFQOKAMTAGOKQ 141
 Db 10 NKQGTGTTLLGGAGGALLGSGQFGKQGVGVGALLGAVLGGQIGAGMDEQDRRLAELT 69

QY 82 LNQSLEKVKAGQVTRWRNPDGNSYSVEPVRYQYKQERRQQYCRFQOKAMTAGOKQ 141
 Db 70 SORALEAPSGSVNWRPNNGHGYTPNKTY-----RNSTGQYCREYTVTVVIGGKQ 124

QY 142 EYGTAC 148
 Db 125 KAYGNAC 131

RESULT 14
 Q8KLU2 PRELIMINARY; PRT; 131 AA.

AC Q8KLU2;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Outer membrane protein (Fragment).
 GN 17 kDa.
 OS Rickettsia helvetica.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsiidae; Rickettsia.
 OX NCBI_TaxID=35789;

Query Match 28.3%; Score 236.5; DB 2; Length 131;
 Best Local Similarity 37.8%; Pred. No. 1.2e-13;
 Matches 48; Conservative 22; Mismatches 48; Indels 9; Gaps 2;

QY 26 SRQEVGAATGAVGVGAGQVTRWRNPDGNSYSVEPVRYQYKQERRQQYCRFQOKAMTAGOKQ 141
 Db 10 NKQGTGTTLLGGAGGALLGSGQFGKQGVGVGALLGAVLGGQIGAGMDEQDRRLAELT 69

QY 82 LNQSLEKVKAGQVTRWRNPDGNSYSVEPVRYQYKQERRQQYCRFQOKAMTAGOKQ 141
 Db 70 SORALEAPSGSVNWRPNNGHGYTPNKTY-----RNSTGQYCREYTVTVVIGGKQ 124

QY 142 EYGTAC 148
 Db 125 KAYGNAC 131

RESULT 14
 Q8KLU2 PRELIMINARY; PRT; 131 AA.

AC Q8KLU2;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Outer membrane protein (Fragment).
 GN 17 kDa.
 OS Rickettsia helvetica.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 OC Rickettsiaceae; Rickettsiidae; Rickettsia.
 OX NCBI_TaxID=35789;

Query Match 28.3%; Score 236.5; DB 2; Length 131;
 Best Local Similarity 37.8%; Pred. No. 1.2e-13;
 Matches 48; Conservative 22; Mismatches 48; Indels 9; Gaps 2;

Db 10 NKQGTGTLGGAGGALLGSQFGKKGQLGVGVGVALIGAVLGGQIGAGMDEQDRRLAELT 69
QY 82 LNSLEKVKAGQVTRWRNPDTGNSYSVEPVRTYQRYNKQERROQYCFEFOCKAWIAGQKQ 141
Db 70 SQALETAPSGSNVWRNPNGNYGYTPNKTY-----RNSTGQYCREYDTTVVIGKQO 124
QY 142 EHYGTAC 148
Db 125 KAYGNAC 131

Search completed: November 5, 2003, 20:15:03
Job time : 25.7824 secs